The Dissertation Committee for Eric Russell Webb
Certifies that this is the approved version of the following dissertation:

The Relational /r/:

Three Case Studies in Rhotic Integrity and Variation

Committee:

____
Jean-Pierre Montreuil, Supervisor

____
Brigitte Bauer

____
David Birdsong

____
Knud LamBrecht

____
Mark R.V. Southern

The Relational /r/:

Three Case Studies in Rhotic Integrity and Variation

by

Eric Russell Webb, M.A., B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May, 2002

Acknowledgements

I begin, as have so many others before me, with a profound expression of gratitude to my dissertation supervisor. Jean-Pierre Montreuil served as the director of this work and deserves much of the credit for its completion. Without his consistent encouragement in moments of frustration and his contagious enthusiasm in moments of progress, this dissertation might never have seen the light of day. It has been an honor to work under his guidance during my time at the University of Texas. I also wish express my sincere gratitude to the other members of my committee: Brigitte Bauer, who has given selflessly of her time and provided useful insight; David Birdsong, who guided me through reviews and whose linguistic expertise have shaped the following pages; Knud Lambrecht, whose love of language and wealth of experience have challenged and sharpened me as a linguist; and Mark Southern, a bountiful source of ideas and a tireless sounding board for new and innovative approaches to linguistic questions.

I have been privileged to work with a number of other fine linguists at the University of Texas and elsewhere. I am indebted to Bjorn Lindblom in particular, whose guidance and input were so crucial in the earlier stages of this work. For their role in my formation as a linguist and their contributions to this work, however indirect, I wish to thank Michelle Biermann Fischer, Carl Blyth, Cécile Canut-Hobe, Mark and Margy Doty, Mans Hagberg, Gudrun Ledegen, Fabrice Leroy, Mark Louden, Jean-Louis Roederer, Gudrun Stevenson, and Edo Velema. Critical financial support for my studies, in general, and for this work, in particular, has come from a number of sources. The Nederlandse Taalunie and the Fonds voor Neerlandistiek in Buitenland contributed generously for travel and research in the Netherlands and Flanders. The Department of French and Italian at the University of Texas and Chair Dina Sherzer have been very helpful in locating additional travel grants. More recently, I was honored with a Continuing Fellowship from the Graduate School of the University of Texas, which permitted me to concentrate on the completion of this dissertation. To all those who have placed their confidence in me, I hereby express my heartfelt thanks.

It is said that any accomplishment lacking an audience of friends is of precious little worth: I am truly blessed by the wonderful individuals who constitute my audience. My fellow graduate students have laughed and studied with me, provided advice and support, and seen me through a process that appeared at times unending. Friends in North America and Europe—too numerous to list here—have provided much needed support and encouragement, as has my family. My sister Heidi, in particular, has shown remarkable steadfastness throughout my many years of graduate work and has never failed to express solidarity with and support for her little brother. To all of these people I express my love, gratitude, and hopes for a future of peace and wisdom.

Last but not least, I would like to thank my partner Ronald Pluth. He taught himself Java and wrote the executables included in this dissertation; more importantly, he tolerated the ups and downs that come with sharing the life of a doctoral candidate—one as difficult as myself—and never failed to say the right thing. Without his love, an achievement of this scale would have no value.

The Relational /r/: Three Case Studies in Rhotic Integrity and Variation

Publication No.______________

Eric Russell Webb, Ph.D.

The University of Texas at Austin, 2002

Supervisor: Jean-Pierre Montreuil

This dissertation articulates a relational phonology of rhotic or r-like sounds from a functionalist perspective. The first chapter notes that no viable, phonetically motivated evidence for the classhood of rhotics has been proposed to date and advances the case for a non-structuralist approach to the question of cross-linguistic rhotic classification, i.e. one that does not begin with an assumed phonological structure, but one where structure is considered a relative, emergent property of linguistic (here, phonological) function. A second chapter presents functionalist theory and its application to phonological classhood, stating that a segment may be defined according to the articulatory and perceptual similarities and differences it presents with regard to a larger, organic whole. The third chapter consists of a phonetic description of the consonant inventories of American English, Amsterdam and Brabant Dutch, and European French, using

native speaker tokens as the foundation for subsequent discussion. The dynamic relations of individual segments within the continuant consonant systems are described in the fourth chapter, consisting of segmental definitions resulting from the similarities and differences present within an organic system. A constraint based, OT inspired framework is employed in discussion and advancement of the segmental integrity (or definition by relational dynamic) of rhotics, in particular. Conceptual mapping of articulatory and perceptual systemic relations is also proposed. Relational definition of individual rhotics leads to the position of these segments' phonological relational integrity or classhood, i.e. the cross-linguistic similarities between the individual relational specification of each of these sounds. A final chapter applies the relational definition of rhotics to instances of variation and shows how a relational definition of the segment contributes to a greater understanding of linguistic behavior.

Table of Contents

INTRODUCTION

The present dissertation provides an innovative concept of the phonological segment, encompassed within a theory I term relational phonology. This work is a preliminary elaboration of relational phonology, as well as an application of this theory to the question of rhotic classhood, i.e. the cross-linguistic unity or disparity of sounds traditionally grouped as possessing an "r-like" quality. My approach to rhotic phonological classhood derives from a functionalist theoretical stance, assuming that the communicative function of language and its subjectivity to biomechanical universals determines particular forms. Furthermore, my application of functionalist principles and implications leads to an innovative definition of the sound segment, whose integrity derives from systemic relation and not from synthetic features.

The five chapters of this dissertation reflect the evolution of relational theory, applied to phonology and to the question of rhotic classhood. Chapter One provides a critical overview of linguistic literature relevant to the question of rhotic classhood and behavior, noting that no viable, phonetically motivated explanation has been put forth that provides convincing evidence for the phonological congruence of different sounds grouped as rhotic. This section also challenges traditional phonological models and serves as a background to the subsequent introduction of functionalist theory. Chapter Two articulates the principles of functionalism as they are used in this dissertation and opposes them to more traditional approaches. Functionalist implications are applied to the linguistic system in a series of articulatory and perceptual constraints, which allow for the generation of theoretical grammars. These theoretical grammars are expressed using violable, systemically defined constraints in an Optimality Theory-inspired, non-generative model. Chapter Three constitutes a departure from theory, examining the gestural and acoustic properties of continuant consonants in three languages having different types of /r/. This chapter lays the groundwork for the advancement of a relational theory of the phonological segment. Integrating the results of this experiment within a functionalist framework, Chapter Four presents the observed systemic interactions and articulates a theory of relational integrity. Applied to rhotics, relational integrity redefines these segments according to system-specific phonological properties and provides for their relational classhood. Finally, Chapter Five examines the implications of a relational approach to the study of rhotic behavior and analyzes three cases of variant rhotic behavior. Using a generative, constraint-based output model, I demonstrate how a relational understanding of the phonological segment is able to describe and explain cases of allophony and variance. This chapter also raises a number of questions as to the scope of phonology and the shortcomings in current phonological approaches to variation, namely those that ignore extra-linguistic factors.

CHAPTER ONE. THE RHOTIC QUESTION

1.1. Introduction

The term *rhotic* makes reference to the set of sounds commonly grouped under a descriptive umbrella as "r-like." In phonetic and phonological literature, the set of these sounds is limited to non-occlusive, non-lateral consonants articulated in the supra-laryngeal buccal cavity. These are produced in no less than four manners (trill, tap/flap, fricative, and approximant), employing at least three articulators (tongue tip, blade, or body) and having one of four articulatory targets (front incisors, alveolar ridge, velum, or uvula). Table 1.1 provides the set of sounds generally considered r-like or rhotic, using symbols and diacritics of the International Phonetic Alphabet (IPA, revised to 1993). Examples of each rhotic are given below the IPA symbol.

It is not by coincidence that most of these sounds are represented in IPA using some variant of R/r in Roman script, thus reinforcing the traditional notion of segmental grouping. Among Indo-European languages—the languages of IPA founders—these sounds constitute important members of consonant inventories. However, the presence of rhotics is not limited to Indo-European. According to Maddieson (1984), 75 percent of the world's languages contain at least one rhotic. Of these, approximately 18 percent have more than one underlying rhotic. It is naturally impossible to state with any empirical accuracy how many such rhotic-endowed languages demonstrate cases of variation, whether allophonic, dialectical or otherwise.

	dental	alveolar	post-alveolar	retroflex	palatal	(velar-) uvular
trill	r̪ Italian [ka r̪o]	r Russian [para]	r̠ Czech [r̠ad]	ɽː Toda [kaɽː]	rʲ Russian [pa rʲ a]	ʀ Swedish (Fin.) [ʀas]
tap or flap	ɾ̪ Italian [kaɾ̪o]	ɾ S. Dutch [ɾam]	ɾ̠ English [waɾ̠]			
fricative						χ̝ ʁ̝ French [paχ̝ti] [iʁ̝e]
approx.		ɹ US Eng. [ɹid]	ɹ̠ US Eng. [ɹ̠ɑd]			χ̞ ʁ̞ French [paχ̞] [paʁ̞]

Table 1.1. Segments traditionally grouped as being r-like or rhotic: IPA (1993) symbols are given, as well as examples for each.

The term "rhotic" is both old and new. Use of the nominalized term "rhotacism" or its equivalent in other languages has been common in phonology throughout the twentieth century, describing among other phenomena the interference of "r-like sounds" on another segment or—such as in the case of many English dialects—the loss of /r/ in word-final environments. According to Walsh Dickey (1997: 70, her p.c.), the first use of the term "rhotic" to encompass a class of sounds was by Dixon (1972) in his description of the consonant inventories of Australian languages. Further use of the term rhotic to describe the (admittedly tentative) class of sounds is seen throughout the latter quarter of the twentieth century, to the extent that rhotic terminology is generally recognized and accepted within the linguistic community.

1.2. Historical Overview

It should come as no surprise that the study of rhotics has long fascinated linguists, whether phoneticians, phonologists, or historians. Having its roots in Europe and, later, North America, the discipline was immediately confronted by a wide variety of "r-like sounds," within the Indo-European family as well as among varieties of each of these. Indeed, some of the earliest linguists—whether *philologues* or *Indogermanisten*—expressed a fascination with rhotics, focusing closely on historical change and the substitution of one type of /r/ for another.

The earliest German philologists contented themselves with observation of different rhotics and their classification. Trautmann (1886) distinguishes between two types of *reine Consonanten* or "pure consonants," *Schleifer* (smooth or polished) and *Klapper* (rolled or flat). Examples of the former are [ʀ] and [ɹ]; [r] and [ɾ] belong to the latter category. While original—and without lasting influence—Trautmann's distinction presents meager, mostly impressionistic evidence for their inclusion in the class of *liquidae* (77-78). Likewise, Sievers (1898) groups the different /r/ of Germanic into one group without distinction of place or manner (87). He notes that rhotics exhibit different behaviors, including resistance to gemination (115-16) and frequency of metathesis (86). Jespersen (1890/1920) acknowledges rhotic variability, discussing place and manner characteristics, as well as phonotactic considerations in differing Germanic languages. He does not, however, provide any extrinsic motivation for this grouping of sounds together but allows their de facto classhood to be assumed (137-145).

Passy (1891) showed little particular interest in rhotics as a category, but rather took the unity of these segments for granted, also grouping these as a category of sounds of unspecified and unmotivated relation. His historical analysis of the change from apical to dorsal articulations merits passing attention, not only for its attention to detail, but for his conclusion that this change might be attributed to productive factors (146-7, 150-3, 156). Chapter Two re-examines such ideas proposed by Passy.

One of the first linguists attempting to explain, rather than merely describe rhotics was Prokosch (1939). His comparison of Germanic dialects rests primarily on the alternation of /r/ and /z/, seen particularly in the history of northern and western Germanic (84-85). To account for this, he reverts to a phonological definition—albeit impressionistically motivated—considering rhotics to be "fundamentally spirants." This classification rests on his observation that "these sounds may be said to occupy a position between consonants and vowels" (45-46). As such, Prokosch was the first to my knowledge to advance a motivated categorization of rhotics—along with other consonants, e.g., laterals—asserting that these are intermediate segments along the consonant-vowel continuum.

Throughout the largest part of the twentieth century, the question of rhotic unity or classhood was ignored or treated summarily. In spite of initial interest in these segments, little work was produced which departed from a speculative or simply descriptive framework. Of the relatively small number of studies, these may be categorized as historical, language-specific, or classificatory. (Obviously, the line between each of these is gradient and many studies may be placed in more than one category). Literature falling into the first group, seeking to explain the diachronic evolution of rhotics, is abundant but of only indirect interest to the subject at hand.

1.3. Language-specific phonetic studies of rhotics

Phonetic investigation of rhotics in specific languages—or their contrast between languages—is especially interesting for the present study. These works comprise a body of literature that provides indirect evidence for rhotics constituting a distinct type or class of sound. I make no pretense of offering an exhaustive survey of all studies of /r/ in this chapter, but rather wish to present a judicious selection of works useful in understanding the discussion evolved in later chapters.

Recasens (1991) investigated the behavior of apico-alveolar trills and taps in Catalan using electropalatogram (EPG) and adjacent vowel, second formant measurements. In response to the question of what phonetic correlates distinguish between the [r] and [ɾ]—rhotics simultaneously present in all Iberian languages—he concludes that these members of the Catalan rhotic set are produced by different, albeit similar gestures. The tap requires less articulatory control and the trill more articulatory constraint. His findings are supported in first part by differences noted in the number of EPG contacts during consonant closure in varied vowel contexts. In the case of trills, less fronting was noted when contiguous to high, front vowels; the reverse is true of taps. Both were shown to be produced in lingual-palatal region subsequent to a back vowel (274). With regard to second vowel formants, Recasens notes that flaps are more subject to vowel-dependant (anticipatory) coarticulation than are trills (277-78).

In a later study, Recasens and Dolors Pallarès (1999) return to the question of distinction between [r] and [ɾ] in Catalan, looking specifically at the coarticulatory effects of these rhotics in VCV sequences. They note that there are significant differences between the flap, where there is little effect on the vowel and where consonantal coarticulation is mostly anticipatory, and the trill, which would seem to override certain lingual gestures of adjacent vowels.

Westbury, Hashi and Lindstrom (1998) offer one of the most comprehensive phonetic studies of any rhotic in their investigation of the lingual properties of the American English /ɹ/. Using x-ray data from 53 subjects, they investigate the shape of the tongue and the position of tongue-to-target contact during the rhotic gesture. Their results underscore the variability of /ɹ/ articulation in English: no gender, age, or physiological pattern could be established to account for the differences seen in subject output. Narayanan, Alwan and Haker (1995) present similar variation data resulting from EPG measurement, albeit in a study involving far fewer subjects. These studies point to a characteristic that has long frustrated linguists attempting to study rhotics, in English or in other language: the wide variability of rhotic production even by speakers of the same language. They conclude that such variability is the result of physiological and production differences. This study provides additional proof as to the triviality of micro distinctions for both the articulation and perception of the American English /ɹ/.

McDonough and Johnson's (1997) study of Tamil liquids offers evidence for the contrastive nature of rhotics and their phonetic interaction with other liquids, specifically laterals. Tamil is a Dravidian language of southern India, in which there is evidence of two rhotics—tap and trill—and two laterals—plain and retroflex. A fifth member of the liquid set has been described as a rhotic, lateral, glide and fricative. Using linguograms and EPG measurements, McDonough and Johnson assert that certain properties of the fifth liquid are both lateral and central. They conclude that this segment is a retroflex—and from that determination, that it is a rhotic—but note that it does not exhibit the phonological behavior of other Tamil contour segments (16, 18). Perceptual evidence from a secondary study indicates that the central rhotic and laterals are often confused, this due to a common lowered third formant in adjacent vowels (20-21). McDonough and Johnson offer compelling evidence for the specification of both rhotics and laterals in the context of a language particularly rich in continuant consonants and suggests that categorization of such segments is dynamically—as opposed to statically or universally—accomplished.

Spajic and Ladefoged (1994) complement the work of McDonough and Johnson in their investigation of the rhotics of Toda, another Dravidian language. According to previous phonological studies, Toda presents the case of no fewer than six trilled rhotics (post-dental, alveolar and retroflex, all of which are either voiced or voiceless). Each is shown to exist in contrastive distribution (37-38). Spajic and Ladefoged quite logically call this extraordinarily large number of supposedly underlying rhotics into question. Of primary concern to the authors are the number of lingua-palatal contacts produced by each member of the set: these points of articulation are shown to vary widely across phonetic and phonotactic environments. Voicing quality is also questioned, as is the effect of adjacent segments on rhotics, such as palatalization and rhotic-to-vowel coarticulatory effects. They conclude that the place distinction between post-dental and alveolar is poorly founded and that the differences between these two rhotics could be attributed to secondary articulations. Spajic and Ladefoged are, however, unsuccessful in determining the nature of the secondary articulation. This apparent failure is interesting inasmuch as it offers evidence of the potential phonetic variation among the set of rhotics. While the IPA gives a finite number of symbols and diacritics (assuming for the moment that rhotics do constitute a motivated class), variation among the set of rhotics is theoretically infinite, as is the case with all segments. Together, these studies of Toda and Tamil call into question the traditional, de facto phonetic grouping of rhotics.

A final study worth mention here concerns the perceptual correlates of rhotics. Shimizu & Dantsusi (1987) focused on the perception in the r-to-l continuum among speakers of English, Spanish, Hindi, Chinese, Korean and Japanese. All but one of these languages has at least one underlying /r/ and a corresponding /l/, Japanese being the notable exception. In all other languages, listener perception of synthesized signals showed a discrete boundary between the two members of the set. This is explained with reference to the underlying inventory status of each segment. In Japanese, however, perception is shown to be gradient: this is attributed to the phonological homogeneity of liquids in the language, where such segments most often surface as [l]. Shimizu and Dantsusi's study is of particular interest to the present dissertation, as it offers evidence for the categorization of [rhotic] and [lateral] being language-specific. When a language lacks distinction between the two consonants, it would also seem to lack distinction between the two categories. The conclusions reached by Shimizu and Dantsusi are supported by West (1997), who studied the coarticulatory properties of English /l/ and /r/ in VCV sequences. Using the progressive replacement of both liquids in actual speech tokens, she notes that listeners are able to correctly "perceive" each segment, using the acoustic clues gleaned from vowel transitions. In opposition to Japanese, English has categorized an /l/ - /r/ distinction: West's study suggests that, in part, perceptual categorization derives from transition. This variable, system-specific notion of categorization will be called upon in later portions of this dissertation.

1.4. Language-specific phonological studies of rhotics

A second body of literature providing a critical framework for the present dissertation is comprised of those studies investigating the phonology of rhotics in particular languages. Most phonological studies attempt to account for rhotic variation and look to environmental and productive constraints to explain situations of allomorphy. Some, such as Tousignant (1987), are quite astute in observing variation and even describing contrasts between different realizations of an assumed underlying rhotic, in this case [ʀ] and [r] of Montreal French. Others are forced to posit more abstracted underlying rhotics, as is the case for Newman (1980), who concludes that while the tap and trill of Hausa are related, the presence of near minimal pairs points to two distinct phonemes.

Elert (1970) introduces one example of dental and retroflex rhotics behaving in accordance with traditional place rules and, by extension, with the obligatory contour principle (OCP) in the dialects of central Sweden.[1] Traditional provisions of OCP prevent adjacent, identical segments from being realized, this in accordance to language-specific rules. In the case of Central Swedish, two possibilities are available for resolution of the sequence of dental rhotic – dental consonant. A rhotic may be either elided—such is the case at morpheme boundaries—or the second consonant may be realized as a retroflex. It is noted in passing, however, that the dialects of southern Sweden and of Finland do not permit consonants to undergo retroflexing, as these dialects use [ʀ] (75). Elert's analysis contributes to the present study by demonstrating that /r/ is not always the phonologically weaker member of a phonotactic grouping, i.e. that rhotics may influence the realization of adjacent segments, just as much as adjacent segments may influence rhotics.

Offering contrast to the phonetic investigation of Recasens (1991), Bonet and Mascaró (1997) present the case for an underlying specification for flaps and

[1] I wish to thank my friend and colleague Mans Hagberg of Gothenburg University for his translation of the relevant chapters this work.

trills in Iberian (here, Spanish, Catalan, and European and Brazilian Portuguese). They critique previous studies, which attempt to explain the presence of trills with reference to an underlying geminate flap (104-5, 119-121). They conclude that each manifestation of flap or trill can be attributed to one of two factors. In clearly strong prosodic positions—absolute word-initial—rhotics will be trilled. Contrarily, in weak positions—final coda or non-initial onset—these will be flapped. The problem of intervocalic contrast is resolved with reference to sonority and underlying specification of "flap" for the gesturally weaker member of the set (114-116). Lipski (1990) unifies surface manifestations of taps and trills using syllabic templates, asserting that the opposition between the two members is the result of template maximization. While Bonet and Mascaró and Lipski do succeed in providing a framework independent of abstracted gemination (never seen in output), their explanation continues the tradition of reliance on input abstraction, in this case specification of the articulatory and/or metrical nature. In their Iberian inventories, [+ flap] and [- flap] members are necessarily present, this with little phonetic evidence for their distinction. In subsequent sections of this dissertation, I call such abstractions into question.

English and its many varieties present a challenge to any researcher interested in rhotics. Studies such as Westbury, Hashi and Lindstrom (1998) highlight the inconstant phonetic nature of the American English rhotic, but such surface variability is but one case of diversity. Not only are there areal contrasts between apical and retroflexed rhotics (as is the case in Scotland and northern England, for example), many varieties of English display as many as three variants of /r/ in allophonic distribution. These are found in southern England and in the coastal regions of New England and the southeastern United States. In such dialects, /r/ may be bunched, retroflexed, or "absorbed" into the tautosyllabic, low vowel. As is often the case, the most interesting literature on English /r/ has been written for pedagogical means, whether to teach non-English speakers the fineries of native-like English rhotic production or as a means of corrective education for foreign language students.

Delattre & Freeman (1968) use X-ray data to contrast variation in American and British pronunciation, concluding that North American English exhibits no fewer than six allophones of /r/, whereas British English shows only two. They note that the instantiation of /r/ in each language depends largely upon phonotactics and environmental constraints: these are attributed to the use of lip rounding and tongue-bunching for rhotics in strong syllabic positions (66-67). Other studies of English have attempted to explain the nature of the so-called "molar-r" (Udall 1958) and of the "intrusive-r" of Received Pronunciation, the form of English used in British erudite circles (Brown 1988). While these do not directly contribute to this dissertation, they do serve to highlight rhotic variability in terms of phonotactic licensing.

Similar to English, rhotics in German vary in both local behavior and geographic distribution. Werler (1980) notes that dialects of Swiss German, considered as a whole, display no fewer than four distinct rhotics. Although each dialect shares at least one rhotic with all others, none utilizes the entire set in its inventory. In a different light, Kohler (1990) shows that, in casual forms of

Standard German at least, /ʀ/ is particularly subject to vocalization, a process similar to that seen in conservative varieties of English. That is, when /r/ is not directly followed by a vowel, it will coalesce and reduce to [ɐ]. This is presented in the general picture of reduction as a phonological phenomenon and is attributed to the perceptual interplay of full and reduced variants, where "the listener decides whether reductions are permissible in transmission of information, and the linguistic group judges the stylistic and social acceptability" (90). Barry (1995a) presents evidence for further post-vocalic variation in German, showing that [ɐ] contrasts with /ə + ʀ/. Barry (1995b) discusses the phonetic distinction between [ɐ] and /ə + ʀ/. This approach harks to Lindblom's articulation of hypo-speech, a theory that is discussed in greater detail in Chapter Five.

1.5. Attempts to unify the rhotics as a phonological class

This brief sketch of phonetic and phonological studies serves two purposes in this dissertation. Firstly, these highlight the articulatory, acoustic, perceptual, and phonological characteristics of rhotics. Secondly, such studies expand the scope of investigation beyond the tradition of "r-like sounds." For the remainder of this work, "rhotic" should be understood as a working term, whose precise nature may be drawn from language-specific or language-contrastive properties, depending on the context of investigation.

Despite the bulk of language-specific work, very little energy has been focused on a definition of [rhotic] as a homo- or heterogeneous category. By this, it is implied that there is one or more unifying features, accounting for both the language-specific and cross-linguistic behavior of /r/. The difficulty inherent to any attempt to unite rhotics under a singular umbrella derives from both their phonetic dissimilarities (varying articulatory and acoustic characteristics) and their particular phonological behavior (allophony, dialectal and historical variation, fortis - lenis oppositions).

Lindau (1985) provides evidence for the lack of a single feature in determining the unity of rhotics. She begins by stating that the phonology of rhotics—their phonotactic behavior and obedience to rules—leads to the conclusion that there is a feature [rhotic]. However, when considered phonetically, she concludes that these sounds do not form a homogenous group (157-58). Her study of ten rhotic allophones in thirteen languages or language-varieties leads to the conclusion that no phonetic correlate exists to explicate the above-mentioned phonological behavior. Previous studies, advocating a lowered third formant, are shown to be valid for consideration of coronal rhotics, though this clearly excludes dorsal rhotics—such as those of French and Southern Swedish—from the category. Such languages' uvular rhotics actually result in raised third formant (163). Parametric relations are, however, shown to exist between each of the rhotics studied. These consist of pulse pattern (for trills), duration of closure, sonorant formants, presence of noise, and distribution of spectral energy or place of articulation (167). Figure 1.1 presents Lindau's representation of this grouping. Stated simply, each member of the set of rhotics is in relation to at least one other: the sum of these relations is what constitutes rhotic unity.

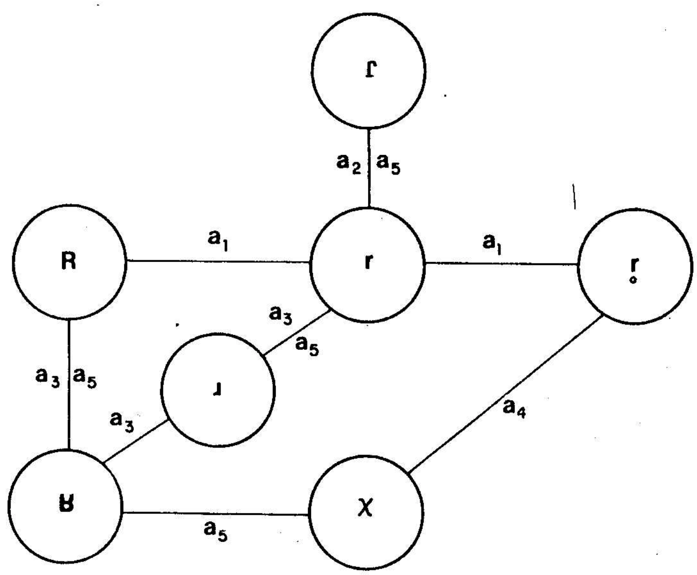


Figure 1.1 (Lindau 1985: 167, her Figure 11.7). Parametric relations among rhotics

Widdison (1997) unites acoustic and articulatory features within Lindau's parametric relation, focusing on the reasons for and directionality of historical rhotic shift. In many languages, such as French and German, apical rhotic articulations ([ɾ] and [r]) gave way to dorsals ([ʀ], [ʁ], and [χ]). Widdison advances two propositions to account for this change. First, he states that vibrants—sounds resulting in a pulse pattern made by rapid opening and closing of the vocal tract—"offer the kind of robust, reliable information that should make them popular for inclusion in sound inventories" (190). This conclusion is based largely upon the quantal theory of speech, as seen in Stevens (1989), wherein speech sounds are said to seek or target areas of acoustic stability. Secondly, Widdison interprets the relative uni-directionality of sound change (i.e., from apical to dorsal, but rarely the reverse) and instances of lenition as indicators of articulatory difficulty. Apicals and trills are said to be more difficult than are dorsals and approximants or taps, respectively. From this, he infers that the apical rhotic [r] is the central—or unifying—member to Lindau's paradigm (190-2). It is not clearly stated in his article whether [r] should be understood as some sort of "core rhotic" or archetype from which all others are derived, nor is it clear to what extent the argued acoustic and articulatory parameters play a role in variation among rhotics.

Ladefoged and Maddieson (1996) offer a succinct overview of many of the above-mentioned studies of /r/, as well as others, contributing a broad, exhaustive discussion of those efforts to provide classification of rhotics. This work provides an important summary of the different members of the rhotic set, highlighting the articulatory and acoustic characteristics of each. However, Ladefoged and Maddieson were not able to further the discussion of rhotic class-hood beyond the conclusions of Lindau. Rather, they reiterate her opinion that rhotics share a "family resemblance" and, among genetically related languages, a common history (244-45).

Walsh Dickey (1997) contributes a phonological analysis of liquids (lateral and rhotic) in one of the most complete treatments of the subject of which I am aware. In first part, she offers cross-linguistic and language-specific evidence for the treatment of rhotics as a polymorphous class. Specifically, she argues that, although there is no one feature uniting all rhotics, Lindau's (1985) observation that each rhotic has something in common with at least one other may be articulated in a way that provides for a multi-dimensional understanding of phonological class-hood. Secondly, Walsh Dickey presents the case for a secondary laminal node.

Walsh Dickey's argument for rhotic unity may be reduced to her position of a secondary laminal node on the feature structure of these segments. This proposal is based on five considerations: positional restrictions, prohibition of rhotic clustering, alternations within the set of rhotics, restriction on coronal rhotics to apical articulations, and resistance to palatalization (1997: 90). While the theoretical and methodological perspective of Walsh Dickey runs contrary to the one outlined in this dissertation, specifically that of functionalism, this work provides a valuable backdrop for a the present approach to rhotics as a distinct phonological class.

By positional restrictions, the first area of her investigation, Walsh Dickey makes reference to the placement impediments in certain languages, such as Australian languages, which prohibit word-initial rhotics (91). The author further makes note of languages with morphological positional restrictions on rhotics, such as Warlpiri, where tautomorphemic sequences of rhotics are equally prohibited. She notes that

> [v]arious phonemic rhotics being selected out of other consonants for phonotactic restrictions is evidence that "rhotichood" can be referred to by the phonology and that "rhotic" is a valid phonological grouping (1997:92).

It is inferred that data from the negatively articulated positional rules seen in the world's languages contribute positively to categorization.

Walsh Dickey's second area of investigation follows from these positional restrictions. These are clustering restrictions, i.e., environmental restrictions, and are presented as either prohibitions against clustering with other rhotics or with non-rhotics. Such phonotactic rules are attributed to the feature geometry of rhotics. Again, the author makes reference to cross-linguistic constraints on rhotic behavior as "evidence that rhotics all have some property that can be referred to by the phonology" (92). In essence, there is little difference between the first two areas of Walsh Dickey's investigation. Phonotactic restrictions placed on rhotics, whether simple negative placement or clustering rules, are considered to be evidence for a common link between each type of rhotic.

A third consideration for class-hood is language-specific rhotic alternation. Walsh Dickey presents evidence from geographically and genetically disparate languages, showing that nearly every language with one rhotic displays some form of allophonic variation (94). In a second stage, she examines rhotic variation in Portuguese, drawing on data from Azevedo (1981), where evidence is presented showing that two underlying rhotic phonemes may alternate between no fewer than nine possible variants, depending upon the dialect in question and when in prosodically neutral positions (Azevedo 1981). She concludes that

> the fact that [ɾ], [ʀ], [r], [ʐr], [ʂ̬], [ʐ], [x], [ʐʁ] and [h] are possible surface variants of *either* rhotic in positions of neutralization is evidence for a class of rhotics. [...] What unites them must be a more abstract quality: membership in the class of rhotics. (1997: 97, italics of the original)

Here, evidence for rhotic unity is posited in rhotic variation.

The most important justification for rhotic class-hood in Walsh Dickey's dissertation, and her final area of concern, is a proposal that rhotic structure consists of a secondary laminal node. As with other considerations, the rhotic structure is attributed to negative evidence, specifically the absence of palatal rhotics in the world's languages and rhotic resistance to palatalization. The first of issue raised in this section is the importance of rhotic place of articulation. This calls into question which part of the tongue is implied in the production of rhotics, as well as the place of articulation of rhotics:

> Rhotics are almost universally restricted to an apical place of articulation. This fact is particularly striking in the languages of Australia, where most manners of articulation contrast apical and laminal coronals. These languages usually have laterals, stops, and nasals which are apical (alveolars and retroflexes) as well as laminal (dentals and palatals). Yet rhotics in Australian are never laminal. Cross-linguistically, dental, alveolar and retroflex rhotics are all apical. And palatal rhotics are conspicuously missing from the inventories of the world. Apical rhotics are abundant, laminal rhotics are avoided. Even rhotics at a dental place of articulation, which is normally laminal, are apical. (98)

Walsh Dickey further presents the case for the statistical insignificance of palatal articulations among the world's languages, based on those languages and transcriptions available in the UPSID database (102), and for rhotics resisting secondary palatalization.

Walsh Dickey's insertion of a secondary laminal node follows from the position articulated by Clemens (1976) and Hume (1992), that palatalized segments have a secondary V-Place, Coronal node. Walsh Dickey gives a geometric model of palatalization, seen in Figure 1.2. She later modifies this model, replacing “Laminal” with “Coronal.” This modification is based on the position that, because the tongue blade is not used in most rhotic production and because most rhotics do not palatalize. The structure of coronal rhotics is presented in Figure 1.3.

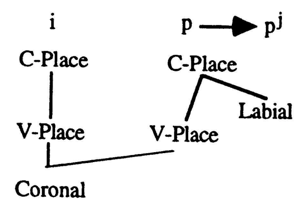


Figure 1.2. (Walsh Dickey 1997: 103-4, her Figure 3.5). Geometric model of palatalization

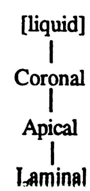


Figure 1.3. (Walsh Dickey 1997: 104, her Figure 3.6). Geometric structure of Coronals

The proposed structure of coronal rhotics is subsequently integrated into an analysis of Carib. Within an Optimality Theory framework, the author demonstrates how her proposition can account for the unusual behavior of rhotics in this language, which do not palatalize in environments where this process is triggered for all other consonants. Summarily, it is assumed that the presence of a secondary laminal node blocks palatalization in a structure-preserving effort

(1997: 106-114). In second part, she admits that palatalization is theoretically possible, allowing for it only in terms of the amount of place structure permitted by a language. Such permission for quaternary structure is mitigated by faithfulness constraints—one for each level—that penalize any change in input place specifications. In her own words, "authorization of rhotic palatalization in a particular language depends on the ranking of the quaternary place constraint; {IDENT Place} must be ranked above {*Place4}" (120).

A final series of considerations, with regard to the place structure of rhotics and the crucial presence of a secondary laminal node, is the treatment of uvular articulations. While Walsh Dickey admits that the inclusion of these segments in her investigation is troubling, she argues for the following structure as means of consistent structural representation, as seen in Figure 1.4 (138).

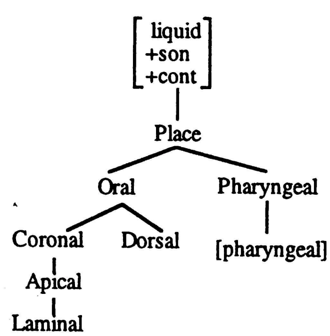


Figure 1.4. (Walsh Dickey 1997: 138, her Figure 3.18) Structural representation of Rhotics

The uvular structure she proposes consists of a branching oral place node, where uvular (and, it is assumed, velar) rhotics are both dorsal and coronal.

Under the coronal sub-node, we find the same latent laminal place structure. Unfortunately, no language-specific examples are presented in this regard.

Walsh Dickey provides a radical departure from purely explicit accounts of rhotic class-hood, i.e. those that look only toward phonetic properties. While a complete critique of her work would be of only marginal use to this dissertation, it is important to note several divergences vis-à-vis the approach used in this dissertation and that of Walsh Dickey. Foremost among these are the differences between structuralist or formalist approaches like her own and those that assume the function of language to be the primary motivation for form. Chapter Two offers critical insight into this question; it is sufficient to state here that the starting point of each work is radically different.

A secondary consideration in regard to her and my understandings of rhotic class-hood is the very definition of what it means to belong to a united or defined class. In order to effectively present her point, Walsh Dickey must offer some means—structural or otherwise—for the union of rhotics, i.e. she must posit their common denominator. As will be shown in following chapters, a functionalist argument for unity need not look to structural similarities, but also to the relational properties of a segment, in opposition to other segments of a same inventory. In essence, the functionalist argument I provide in this dissertation states, much like Passy (1891) whose work has been one of its greatest influences, that a segment may be defined by its systemic differences and similarities as much by its tactile reality. The functionalist approach outlined in this work begins by assuming that rhotics are not an a priori class, but an emergent class, whose unity derives from the systemic relations of rhotics within the context of linguistic function.[2]

1.6. The need for a relational phonology of rhotics

Like Walsh Dickey, I assert that it is possible to provide for the phonological unity of rhotics and, in concert with Lindau and Widdison, I assume that this unity must look across languages and to the congruous properties of different rhotics for its substance and advancement. Unlike Walsh Dickey, I do not assume that form predicates function or that the feature of one language or language form, however motivated, is necessarily available to another. In opposition to Lindau and Widdison, I state that a segment must first be defined with regard to its local, organic context and only secondarily compared to other, similar segments defined according to and within their respective contexts.

The thesis of this work is that rhotics constitute a particular phonological grouping (or class, to use traditional terminology) by virtue of congruous systemic specificity, that is the ensemble of similarities and differences that allow for each rhotic to be considered a distinct segment in its respective phonological system. As such, rhotic unity is a relational construct, deriving from the interaction of one segment to another. The idea of constructing a phonological grammar using an system of contrasts and similarities is hardly new to this work, nor is the idea that

[2] Use of the term "emergent" makes reference to the classification of sounds from the bias of the use (productive and perceptual) of such sounds; emergent classification issues from the linguistic system. An emergent class may thus be opposed to an ad hoc class (see, notably, Diehl and Lindblom 2000).

the analysis of phonemes and of phonological behavior must be undertaken from the bias of the holistic language system (see, for example, Saussure 1916/1967: 164). Lindblom (1990, 1983, and 1972) and Flemming (2001, 1995) provide in their Dispersion Theory (DT) a construct within which the relative harmony of system members is considered. Schwartz et al. (1997a and 1997b) use a modified form of DT, Dispersion-Focalization Theory, to generate hypothetical vowel inventories. And Padgett (2002, 2001) and Ni Chiosain and Padgett (forthcoming) specifically use such principles to analyze issues of historical phonological change. The originality of the present work consists of two crucial additions to the above, namely that phonemes are deduced from a relational or systemic perspective and that relation of these to the linguistic whole is necessarily made prior to any discussion or analysis of variation. This deduction is heavily dependent—in its early stages at the very least—on phonetic analysis, leading to phonemic definitions that are highly grounded in articulatory and perceptual reality. The present work is also innovative in as much as it addresses systemic relations, having consonant inventories as its focus, and compares systemically defined segments in one language to another as a means of promoting not classhood, as this term is generally understood, but relational similarity or congruence. The term “relational phonology” thus derives from the work undertaken in this dissertation, as well as antecedent works, such as those mentioned above.

In order to advance a theory of relational phonology, I provide two foundations essential to the relational treatment of rhotics or any other segment.

The first of these grounds my innovative notion of phonological relations in functionalism, an approach that stands in opposition to formalist or structure-primary conceptions of the phonological segment. Functionalism assumes that linguistic form derives from the particular communicative function of language and from its subjectivity to human biomechanical principles. All rules employed in discussion of rhotic variation—articulated as violable constraints and presented in an OT-inspired framework—are borne from these supralinguistic universals, rather than from rules posited from observations of particular linguistic form.

The second step involves the phonetic description of those sound segments with which the rhotics are in relation. In this dissertation, phonetic investigation serves to provide proper grounding of phonological discussion and categorization; while not exhaustive, the phonetic sketch of different continuant consonants and of different rhotics permits the postulation of organic, system-emergent features to be used in phonological discussion. These features contrast with those of structural phonology, mentioned above, which begin with a predetermined set of features that are assumed available to all languages.

I focus on three different languages in my evolution of a relational theory of phonological segments, in general, and of rhotics in particular. These are American English, Dutch (Netherlandic), and European French. The choice of the three subject languages was made in part for practical reasons, but also as a means of focusing discussion in the present work on a limited data set. Future investigation of other rhotics and of other phonological systems can only further and reinforce the theoretical groundwork I lay in this dissertation.

CHAPTER TWO. FUNCTIONALISM AND RELATIONAL GRAMMARS

2.1. Introduction

The previous chapter presents an overview of the study of rhotics, noting the lack of a convincing story that accounts for the apparent unity of these sounds. Those analyses that have been put forth offer either only indirectly sustainable motivation for the assumed unity—such is the case of Lindau (1985)—or rely on structural abstraction—as in Walsh Dickey (1997). These inductive approaches are adequate for the description of an assumed class, but do not explain such classhood, nor do any attempted explanations look past specific instances and contribute to a larger understanding of language, i.e. have any heuristic power. Such shortcomings testify to the limitations of traditional phonological analysis. The use of ad hoc formalism results in the position of structures that do not derive from *le langage*—the human communicative propensity and capacity to achieve communication through articulation and perception. Rather, such structures are born from *la langue*—specific manifestations of the human language capacity, wherein countless linguistic and non-linguistic variables contribute to observable reality.

This chapter outlines a theory of language that is later used to describe rhotics and rhotic behavior, looking specifically at how these segments are related to others within specific phonological systems. In first part, I summarize the history and principles of functionalism, the theoretical perspective underlying my work. These are contrasted with approaches discussed in Chapter One. Secondarily, I provide a series of functionalist principles and implications: the former are global in nature, whereas the latter apply specifically to human language and are applied to segments resulting from speech acts. Regarding the question at hand, that is a relational description of sound inventories and of rhotics place within these, I describe the interaction of two linguistic drives, perception and articulation. Functionalist principles and implications are expressed locally in an OT-compatible format, consisting of ranked, violable constraints. These functionally grounded constraints provide example descriptions of theoretical inventory grammars.[3]

2.2. Functionalism: A Historical Overview

Passy (1891) was one of the first linguists to assert that human speech is organized in accordance with larger, language-external principles. His foundational work on the history of languages and on sound change cannot be directly related to the subject of the present study. However, Passy's perspective on the catalysts for change may be reconciled with the theoretical bent of this dissertation, inasmuch as he attempts to look beyond the surface form of linguistic

[3] The reader will note the lack of discussion of generative phonology in this (and later sections of the dissertation). By this, I make no claim that functionalist or relational treatment of phonological variation is non-generative. The reader should further be aware that, in opposing functionalism to formalism and structuralism, I in no way assume that structure and form (or formalism) have no place in phonological analysis. Rather, I assume that any structure or form results from and must be motivated with reference to linguistic function.

phenomena, i.e. the changes themselves, to the dynamic underlying and motivating such change. He notes specifically that

> des tendances phonétiques que nous avons constatées, se dégagent bien nettement deux principes fondamentaux:
>
> 1. Le langage tend constamment à se débarrasser de ce qui est superflu.
>
> 2. Le langage tend constamment à mettre en relief ce qui est nécessaire. (227, numbering of the original)

These two principles are further expounded to include all instances of *économie* and *emphase* (228), from which all changes are purported to derive. Passy further states that

> [a]u fond, les deux principes, d'économie et d'emphase, dérivent d'une seule et même cause, *celle qui motive l'existence du langage lui-même: la nécessité de se faire comprendre.* ***On parle pour être compris****, et rien que pour être compris* [...]. (229; italics mine, bold type of the original)

For Passy, the human propensity to use language, and the form of that language, is subordinate to the relatively simple premise that humans speak in order to be understood, and for no other fundamental reason than to be understood, whence derives the principle of economical and clear communication. Clarity and economy are not, for Passy at least, terms attributed to form, but to gestures and to acoustic signals.

The primacy of communicative function over form was not opposed, but largely ignored by Saussure. Although not directed against ideas such as those generated by Passy, this work resulted in the strict distinction between form and function, wherein form was assumed primary to linguistic production. This foundational work did not deny the importance of function—indeed, Saussure acknowledged the importance of systemic distinction and of linguistic function. Rather, Saussure's approach to questions of phonological description and classification specifically relegate phonetics and the study of speech production to a secondary or post-phonological role (1916/1967: 63-64).

Jakobson and the Prague School contributed to the traditional primacy of form over function, more precisely of phonology over phonetics. For Jakobson (1963), speech could be reduced to a series of distinctive features, each of which represented a concrete, formal property and from which surface structure was evidenced (104-106). By making a rigorous distinction between phonology and phonetics (or "phonemetrics"), Jakobson emphasizes that phonemes possess internal structure (108), i.e., phonemes are the most basic building block of spoken language, from which physical sound is merely a secondary manifestation. From this, it is implied that a phoneme, while not necessarily present in every language, can be reduced to a combination of universally available features. These features give the phoneme potential to be realized as sound, observed as tactile reality.

Jakobson's perspective constitutes an important page in the history of modern linguistics. During the early and middle part of the twentieth century, the distinction between underlying form and surface output served as a catalyst for advances in phonological theory, without which our current understanding of phonological and phonetic form would be impoverished. However, by making secondary the role that linguistic function might have in the eventual definition of such form, structural approaches to phonology soon found themselves reliant on more and more abstraction.

Martinet (1955) represents a return to the externally motivated ideals laid out by Passy, again in a work focusing on historical change. One of the most important contributions of Martinet's work to modern linguistics was his strict opposition of function and structure, here related as form. In Martinet's view, function is a distinctive communicative operation, one deriving from the act of speech and not from the content thereof.

> Le terme [fonction] est pris ici avec sa valeur la plus commune d'activité propre, de rôle bien adapté à la nature du sujet agissant [...] avec toutefois cette réserve qu'en linguistique, l'activité dérive de l'homme qui parle et non des unités phoniques ou de sens à qui l'on prête des fonctions. (39)

To this, Martinet opposes the structuralist tradition, whose approach is rejected as being too synchronically oriented (63). Further critiqued is the penchant for binary oppositions (72-74) and the "integrationist" presuppositions of his contemporaries (88-89).

Rather than repose his analysis on pre-conceived structures, Martinet asserts articulatory economy as a primary catalyst for sound change. This is linked to the pioneering biomechanical and cognitive work of Zipf, whose "principle of the least effort" served as a foundation to early functionalist theory (1949: 94). While the heart of this original discussion is of little direct interest to this dissertation, the application of linguistically external motivations for sound change contributes to the functionalist story. Rather than assume change should be directly attributed to local, internal changes, Martinet provides a framework wherein the mechanisms of language itself are called upon in order to explain the variable nature of forms. These concern markedness, stability, and articulatory mechanisms.

Boersma (1998) constitutes one of the first works attempting to formalize and further previously articulated functionalist principles, such as those of Passy, and integrate them into a phonology capable of treating synchronic and diachronic phenomena. Here, the distinction between form and function is much more radical and complete. As with the present work, which owes much to him in way of inspiration and methodology, Boersma's guiding principle is a distinction between articulatory and perceptual considerations. These are articulated as perception and production grammars, distinct yet related components of human speech (1998: 145-47, 269-70). He further departs from formalist tradition by asserting that features are learned, not shared, by speakers. From the learned features and the competing drives of a larger grammar, the need for and accomplishment of categorization arises (179). Categorization, or the regularization of differences and similarities, is viewed as an emergent phenomenon, as are the constraints that accompany them. Under Boersma's approach, sounds do not possess internal geometry or features, but are given these during language acquisition. Any proposed form or structure of a given segment is therefore an emergent property of its use, and not intrinsic to its production or perception. Furthermore, such categorization is relative to each language system (161-171, 276-94; see also 4.4 of this dissertation).

Given the scope of the question at hand and its departure from tradition, it is useful to develop a series of guiding principles for use in further discussion. While I do not pretend that the following are complete, they do provide a firm foundation upon which to further discussion and in which a theory of relational phonology is grounded.

2.3. Functionalist Principles[4]

This section outlines a series of principles that, for the remainder of this paper, I consider to be the guiding principles of investigation into the rhotic question and the nature of systemic linguistic relations. These assume that language is a biological, as well as psychological phenomenon and that linguistic production, in its largest sense, is subject to the same biomechanical restrictions and constraints as all other human activities. The below principles are applied later to the question of human speech production in a series of implications for this specific type of communicative activity.

The first and most basic tenet of functionalism is that humans share a basic need to communicate. In the most general terms, I attribute this need to the biological fact that we are born of humans and to the sociological observation that we mature alongside and spend our lives within collective, human units. Such communication requires understanding on the part of those humans to whom it is destined. This statement of communicative need is neither new, nor

[4] For the remainder of this work, "functionalism" and "functionalist" should be understood as terms deriving from the works of Martinet and Boersma, i.e. grounded in European functionalist tradition.

revolutionary, and echoes the works of too many predecessors to name. This is given as Principle 1, below.

> **Principle 1**: Human interaction requires communication; from communication arises the need to be understood.

Beginning with this most simple premise, a series of secondary principles are derived. Namely, that a communicative act requires three elements: a message (the communicated lexical information), a means of transmitting that message (articulation), and a means of receiving the message (perception).[5] These are given below.

> **Principle 1.a**: Communication requires specification of communicated information.

> **Principle 1.b**: Communication requires articulation—the activation of some gestural mechanism—to be realized.

> **Principle 1.c**: Communication requires the perception of articulation.

> **Principle 1.bis**: Effective communication requires the perception of articulation be correctly corresponded to the originally specified information.

Having provided principles concerning the most basic nature of human communication, I now approach the question of how this is accomplished. I assert that each of the above-mentioned components of communication is subject to biological constraints. This assertion derives from the observation that, regardless of what is intended by such terms as "articulation," "perception," and "specification," these are accomplished by biological entities, namely Homo

---

[5] Note the parallel between my description of communication and that of Saussure (1917/1967), which depends on the same message, but refers to actors instead of actions. My articulation of Principles 1.a through c is inspired by Boersma (1998, 1-5).

sapiens. Whatever abstraction one may posit in regards to the content of communication, there can be no mistake as to the subordination of this content to the biology of bone and muscle and to the physics of movement, acoustics, and audition. Each of the three components of communication is independent, i.e., each is subject to particular constraints, which may or may not interact with other components. Principle 2 is provided below.

> **Principle 2**: Specification, articulation, and perception are independent, biologically motivated variables.[6]

The foundation to this perspective rests on specification, which I assert is the genesis of any communicative message. I suppose that this component, in regard to the language phenomenon, emerges from the conscience of human communicators. This impetus is generically referred to as the lexicon, as the following:

> **Principle 3 (local)**: Specification is a lexical domain.

Due to the limited scope of this dissertation, specification is ignored, if not completely taken for granted, for the remaining discussion.

Of more importance to the question at hand is the interplay of articulation and perception. Zipf (1949) was one of the first human scientists to focus on the theretofore assumed principle of least effort (56, 132-133), proposing that the most favored of human activities make use of low-cost energy expenditures. In linguistics, this principle has been applied to language change (Lindblom et al. 1995), to variation (the "Hyper-Hypo" theory, Lindblom 1990), to vowel

---

[6] Cf. Boersma (1998: 27) for a slightly different articulation of the specification – articulation – perception triad.

inventories (Schwartz et al. 1997a, 1997b; Lindblom 1983) and to categorization (Diehl and Lindblom 2000; Boersma 1998; Padgett 2002, 2001). In each of these works, ease of articulation plays a primary role in determining which gestures and coordination thereof are allowed. Following from these observations, and from the general observation that humans do not tend to use more effort than necessary for biomechanical tasks, I posit that energy expenditure is universally avoided, as in Principle 4.

> **Principle 4**: Humans will, ceteris paribus, avoid the expenditure of energy.

This principle directly correlates to Boersma's *EFFORT constraint, which states that "we are too lazy to spend any positive amount of effort" (1998: 149). Principle 4 does not deny that humans put forth effort—clearly, this cannot be supported. It asserts however that humans do not do more than the strict minimum in a given, organized gestural behavior. All other things being equal, the strict minimum is nothing.

It follows from these observations on the nature of human action and reaction that perception is obedient to similar, biologically motivated principles. Whereas articulation depends on movement of articulators, perception relies exclusively on neural stimulation, i.e., the activation of certain neural networks and the non-activation of others. It is the differentiation between excited and non-excited networks that allows for distinction between different articulations. It may be inferred that distinction is the outcome of perception: for perception to have significance beyond mere audition or neural stimulation, a human perceiver must distinguish between different stimuli. Here, both positive and negative statements are necessary. Humans do distinguish between different stimuli; however, they also confuse similar stimuli (i.e., stimuli that are not critically different). The tension established leads to the following position:

> **Principle 5**: Humans will, ceteris paribus, distinguish between items that are dissimilar and confuse items that are alike.

This approach differs somewhat, in its articulation at the very least, from Boersma, who formalizes perception as a *CONFUSION constraint, stating, "we are too petty to allow any positive amount of ***confusion***" (1998: 173, bold italics of the original). I further the positive inference of Principle 5, emphasizing the tendency to maximize distinction, and therefore difference, when items are dissimilar: this definition owes much to the dispersion (and, later, focalization) theories outlined in Lindblom (1983), Liljencrants and Lindblom (1972), Flemming (2001, 1995) and Schwartz et al. (1997a).

> **Principle 5.a**: Humans best perceive items that are dispersed, i.e., that are acoustically as different as possible.

It follows logically from this consideration that a negative statement also be provided. Under this perspective, in spite of a penchant for distinction, the economy of Principle 4 militates against too much distinction and, by extension, too much effort.

> **Principle 5.b**: Humans make no more distinction than absolutely necessary between perceptual tokens.

Principle 5.b results in a de facto tension that cannot be fully resolved. On the one hand, distinctive effort is fundamental to human activity. On the other, avoidance of effort is equally endemic. Furthermore, numerous studies attest to the perceptual ability to repair or overcome confusion present in the acoustic signal.[7]

Neither other functionalists nor I assume that Principles 1 through 5 are the catalyst for communication; furthermore, I do not propose that the universals underlying these directly or indirectly motivate communication. The need to make oneself known and to know another is a phenomenon whose fundamental nature escapes the linguist, as it cannot be directly attributed to the reality that is the focus of our discipline. One must, for the moment, assume that communication simply is, and that it is, as stated above, endemic to human existence. To this, Principle 6 is added, defining the nature of biomechanical universals in the manifestation of this phenomenon.

> **Principle 6**: Communication is neither determined nor motivated by Principles 1 through 5; rather, communication is so governed (i.e. communication is subject to limitations implied by these Principles). It is from the positive implication manifest in Principles 1 – 5 that possible communication arises.

Later sections of this chapter demonstrate how these principles may be applied to the description of language inventories and to the evaluation of variation within these constructs.

---

[7] Discussion of psychoauditory and neurophysiological mechanisms is treated summarily in this and later chapters. In presenting perceptual data, I make no attempt to explain or describe the manner by which listeners process acoustic material; rather, I assume quite simply that perception happens (see Holt & Kluender 2000, Lotto and Kluender 1998, and Kingston 1994 for discussion of general auditory and perceptual processes).

2.4. Functionalist Implications

Given the preceding principles, what can then be deduced in the way of possible human communicative activity, more specifically the act of speaking? The following statements synthesize many of the works mentioned in 2.1 through 2.3 and provide a series of informal, working implications that will serve as a framework for discussion of the articulatory and perceptual components of human speech activity in this dissertation:

> **Implication 1**: Humans speak in order to communicate, i.e., to communicate a message (in order to be understood, in Passy's terms).
>
> **Implication 2**: Human speech activity is governed by biomechanical universals.
>
> **Implication 3**: Humans will not, ceteris paribus, undertake difficult gestures in the production of speech events.
>
> **Implication 4**: Humans perceive best, ceteris paribus, sounds resulting from a speech event that are the most different (one from another).
>
> **Implication 5**: Humans do not like to make more distinction than necessary.
>
> **Implication 6**: The best phonic or speech events are ones that result in the least amount of energy expenditure, but which allow for the maximization of distinction between one event, or portion of that event, and all other events, or portions of the same event.
>
> **Implication 7**: Speech events are categorized, i.e., the humans involved in them seek out regularities and attach to these regularities values that may be applied to further speech acts.

For the remainder of this dissertation, I focus on one particular type of speech event, namely that of the phoneme and its allophonic or situational variants. A lengthy tradition defines the phoneme as a unit of the larger phonological grammar that is susceptible to effect meaning; I see no need to change or modify this understanding.

Building from these working implications and returning to the question of rhotic phonological unity, a different approach to the analysis of phonological segments may be put forth. Rather than assume, as do theories of phonology where form is assumed primary to function, that a segment may be understood in terms of universal feature theory or analyzed with a given set of internal structures, the phonology evolved in this dissertation presumes that any phonological segment must first be understood as a product of human communication.

Essential to the methodology developed in this dissertation is the distinction between the two "drives" of human linguistic production: articulatory and perceptual.[8] The use of these terms, and much of my understanding of them, follows the methodological distinction outlined by Boersma (1998). It is useful to speak of drives in terms of the constraints—or violable rules—inherent to them. While there are similarities in the statement of such constraints, their implication for linguistic production differs greatly between the two drives (Boersma 1998: 29-30; Russell Webb, 2002). In the following section, I provide a general discussion of each drive, highlighting some of the more basic constraints that interact within them.

---

[8] One should note that no claim is made or implied by the distinction of articulatory and perceptual drives to the non-relation of these, nor does the division of labor between the drives constitute an attempt to answer the auditorist versus gesturalist debate in phonetics. Kingston (1992) discusses the interaction of articulatory and auditory mechanisms, noting specifically the accommodations made by each in the pursuit of the best possible perceptual signal.

2.4.1. THE ARTICULATORY DRIVE

Boersma makes preliminary reference to the articulatory drive as "articulatory implementation," that is the physical movement of articulators in precise manners, at precise times and in precise combinations away from a neutral state or rest position. These movements have as their goal and result in acoustic events (1998: 9). The notion of "drive"—in both articulatory and perceptual contexts—is explicitly joined to the principles of 2.3; in order to achieve effective communication, each of the two systems (one productive and the other receptive) is involved. Studies attempting to quantify articulations in terms of the relative ease or difficulty of production suffer from the lack of precise definitions for "difficult" and "easy." One earliest use of articulatory effort as a crucial factor in language form is Zipf (1949), who relies on more intuition than on measurable reality. More recently, Lindblom (1990, 1986) provides a synthesis of linguistic and non-linguistic studies, showing how linguistic behavior is mirrored by other activities. Among these, a parallel is drawn between positive exertion of force for such tasks as running and bending, and the movement of articulators in the production of acoustic signals. Boersma (1998) provides a more limited approach, choosing to focus on phenomena that are easily categorized. He notes that—in terms of linguistic articulation—duration, speed, and distal displacement of articulators are factors in the relative ease or difficulty of a particular articulation. I consider these terms as points on a continuum, rather than binary poles. For example, one may only speak of duration if there is movement and may only speak of longer or shorter duration if there is more than one movement within a given system of categorized displacements.

Given that gestures are indeed made, it is possible to note a number of regularities among them. Firstly, simple (i.e. non-complex) movements tend to predominate in the production of discrete units of sound. This reflects Lindblom's principle of "do no more than is necessary" (1990) and is a correlation to Boersma's *GESTURE constraint (1998: 152-3).

Within a system, simple movements and less complex combinations of movements also tend to be reused, creating a relatively small inventory of gestures and gestural combinations, when compared to the theoretically infinite number of gestural possibilities (Maddieson 1995). I assume, much like Boersma, that this minimization of exertion is a result of a broader law governing human activity. The optimization of the human biomechanical mechanism provides that, all other things being equal, we will not do anything, i.e. we will not expend gestural energy. However, should we do something and, later, be required to perform this task again, we will accomplish this in the most efficient manner, repeating the prior gesture. These considerations are captured by the following constraints:

> *EFFORT: We are too lazy to do anything (Boersma 1998: 149, adapted)
>
> REUSE: reuse articulatory gestures (may be stated as *Innovate: do not innovate, i.e. make use of non-native gestures)

REUSE, as defined below, minimizes the burden of categorization within a gestural system. This constraint may be defined locally (within a subsystem) or globally (within the larger system of a language). It is clear that *EFFORT is dominated in every grammar: humans do indeed engage in articulatory activities. However, when doing so, efficiency overrides complexity.

Assuming that gestures are indeed made (i.e. that *EFFORT constraints are dominated in every language and by every inventory), *EFFORT constraints must be mitigated in order to allow for the production of gestures in an efficient manner: this is formulated in REUSE. Here, reference is made to the pronounced propensity to reuse similar gestures in different coordinated combinations. These gestures, such as tongue tip-to-alveolar ridge and labial closure figure prominently in phonological inventories. For example, the tongue moves from a neutral position (where tongue muscles are not contracted) to the alveolar ridge by contraction of tongue muscles in order to produce the stop [d] or [t]. Similar articulator movements are involved in the production of [ð], [l] and [ɾ].

This movement may be considered from three different perspectives: the path, speed, and duration of movement away from neutral. Following Boersma (1998), I advance the following local *EFFORT articulatory constraints:

> *DISTANCE: an articulator does not make positive distal movements from neutral position (Boersma 1998: 150)[9]
>
> *SPEED: an articulator does not complete its movement in any positive duration (=*FAST of Boersma 1998: 151)
>
> *HOLD: an articulator is not held in non-neutral position for any positive amount of time (Boersma 1998: 150)

The ranking of effort in each of the three categories allows assignment of the following fixed rankings (applicable to all *Effort constraints):

---

[9] The reader will note the continued use of "positive" in constraints and in phonetic descriptions (e.g. positive movement). This follows convention in the relevant literature and should not be understood as being opposed to or contrastive with any hypothetical negative movement.

*EFFORT n(x) >> *EFFORT n(y) *iff* x > y

“It is worse to expend effort n(x) than n(y) if and only if (x) incurs positively more effort than (y)”

Note that the specification of each *EFFORT constraint can be phonetic, e.g., *HOLD (2ms) or “Do not hold a gesture for 2 milliseconds,” or phonological. Phonological specification provides for relative ranking, e.g. * HOLD (2) “do not hold a gesture for duration of 2, where 2 = a duration of more duration than 1 and less than 3.” Regardless of which approach is used, universal ranking provides for a theoretically motivated, constraint-based approach to the categorization of articulatory segments.

Table 2.1 provides a generic model of articulatory constraint interaction for gestures A and B, where the articulator of A is held for a positively longer duration than the articulator of segment B. The qualitatively ranked *EFFORT constraint *HOLD provides that B is a better candidate than A. While each segment results in a violation of *HOLD, ranking provides for a “less bad” outcome: the result of constraint interaction is that a relatively shorter movement is less bad than a longer one and it is selected as the optimal candidate.

{A, B}	* HOLD (2)	* HOLD (1)
A	*	*
B		*

Table 2.1. An Example of Articulatory Constraint Interaction

Note that no selection or generation of an optimal candidate is implied in this picture of constraint interaction. The goal of describing segments with violable, articulatory constraints is the eventual description of their interaction within a dynamic, multiplex whole, i.e. a linguistic system.

2.4.2. The Perceptual Drive

The perceptual drive describes the physical and cognitive propensity to turn acoustic cues resultant from articulatory activity into units of meaning, i.e. to associate a physical sound with a categorized sound. Boersma refers to the perceptual drive as "perceptual specification" (1998: 9), providing for the mapping of physical events to acoustic—and eventually auditory and cognitive—regularities. In spoken human language, the perceptual drive is directly dependent on the acoustic signal that results from articulation. In this dissertation, reference to the perceptual drive and to perceptual output should be understood as being the acoustic result of an articulatory act.

Numerous studies have investigated the phenomena of acoustic regularities and irregularities in spoken language. Each has come to a fairly intuitive conclusion: given that a theoretically infinite number of acoustic signals are possible in human speech production, the number actually used in the world's languages is surprisingly restrained. Stevens (1989) shows that sound systems—specifically, vowel inventories—favor regions of acoustic stability. These regions allow for a greater area of "confusability," where the link between acoustic sound and categorized sound is more tolerant of microvariation and allows for easier repair of one to the other. Areas that are more stable are exploited to a greater degree in languages, whereas those that are less stable tend to be avoided.

Similarly, inventory dispersion theories have highlighted the need for perceptual salience, as defined by maximal perceptual discrimination. Lindblom (1983) and Liljencrants and Lindblom (1972) provide that vowels in an inventory are maximally dispersed. This notion is reflected in DT, as provided by Flemming (2001, 1995) and used in phonological analysis of palatalization by Padgett (2002, 2001). Such theories consider maximal or maximized perceptual dispersion of elements in a given system to be the ideal situation for the human perceptual mechanism.

Quite naturally, dispersion as provided by Flemming, Padgett, and Lindblom is not absolute; no phonological inventory makes use of maximal dispersion, as this would result in unusually and superficially complex systems. Schwartz et al. (1997a, 1997b) adds to DT the notion of focalization of acoustic space. Not only must segments be dispersed, i.e. differentiated one from another, but the total use of acoustic space must also be focused. By focus, reference is made to the relative restriction of the entire acoustico-perceptual space. Rather than use the total range of acoustics available, human speakers confine their production to a restrained acoustic space.

Discussion of the perceptual drive in this paper does not make reference to neurological perceptual mechanisms, but rather to the raw ingredients—i.e. acoustic signals—that are susceptible to be perceived by the undamaged human system. In essence, this approach asks, all other things being equal, which acoustic material provides the best possible input to the perceptual drive? Relevant to the conceptualization of the perceptual drive are Principles 1.c., 1.bis, 4, 5, 5.a., and 5.b. From these a two-part tension is evolved: on the one hand, perception will be easiest when segmental acoustic distinctions are maximal; on the other hand, superficially maximal distinctions will result in inefficient or "high cost" perceptual activities. The optimal inventory therefore balances these competing needs: it promotes maximal ease of perception but penalizes superfluities or exaggerations of acoustic distinction. Such penalization does not imply a given distinction will never occur, but that, in perceptual phonological grammar, segments that contain such superfluous contrasts will be unstable and be avoided.

As with the constraints provided in 2.4.1, perceptual constraints result from global optimization of human biomechanical activity. Humans perceive acoustic signals without significant effort: the hearing of signals is, for the undamaged system, a reflexive activity. Optimization states that, all other things being equal, humans will do nothing that results in the positive expenditure of energy: energy is not spent unless it is necessary to do so. Expressed as formal constraints, optimization provides for the following:

> PERCEIVE: perceive signals in acoustic space
>
> *CONFUSION: We are too petty to allow for confusion; by extension, we are too petty to allow for confusion in categorization (Boersma 1998: 173)

As raw formalizations, *EFFORT and PERCEIVE represent the building blocks of the perceptual drive; it is from their interaction that the perceptual drive frames linguistic output. Out of a number of more specific perceptual constraints relevant to the relative definition of sound segments, I have chosen to concentrate on two, first introduced by Flemming (1995) as components of DT. Here, emphasis is placed on the inventory's minimization of distance between segments—favoring the minimization of their confusion—and the maintenance of distances between segments (Boersma 1998: 360).

Most previous studies of inventories have focused on vowels, segments exhibiting more consistent and salient acoustic characteristics than consonants, which are recognized by the type and not only the place of acoustic energy (Flemming 1995; Schwartz et al. 1997a, 1997b; Liljencrants and Lindblom 1972). Several more recent studies have looked at liquids, specifically [l] (Padgett 2002, 2001; Lotto and Kluender 1998). Because vowels and [l] present a distinct set of acoustic characteristics, a certain amount of manipulation of both Flemming and Boersma's original constraints is necessary. The definitions given here are my own, adapted as necessary to this paper.

> *MINDIST *n* (x): there should not be a more than *n* quantitative distances between a segment x and all other segments

> MAINTAIN *n* (x): maintain at least *n* qualitative contrast(s) between segment x and all other segments

Flemming's original constraints have been modified in other recent works, notably Padgett (2002, 2001) and Ni Chiosian and Padgett (forthcoming). In later works, *MINDIST and MAINTAIN are replaced by SPACE and *MERGE, respectively. For the context of discussion of systemic relations in this and subsequent chapters, the original DT terminology have been modified to a point where such titles are of only secondary importance.

Theoretically, an infinite number of perceptual distances and contrasts are available, just as an infinite number of articulations and resultant sounds are possible. In any theoretical discussion, these distances and contrasts are necessarily relational and not absolute: for example, a distance of 2 is defined as greater than 1 and less than three, within a tripartite opposition. Such limitation is supported in both the articulatory and perceptual literature. Lindblom (1990 and 1983), Lindblom et al. (1995), and Carré and Divenyi (2000) make explicit reference to the biomechanical system's ability to reduce overall articulatory effort through the simplification of gestures and gestural combinations. Likewise, the work of Repp (1988, 1986), Repp and Liberman (1984) and Repp and Mann (1981) demonstrate that the human perceptual mechanism does not seek discrete distinction, but rather minimal discriminability, and that covariance in the auditory signal is regularly resolved.

The perceptual constraints *MINDIST and MAINTAIN differ in their implication, reflecting the tension between maximum differentiation and efficient use of acoustic space. When ranked, the former constraint prefers that segments continually reuse the same acoustic space, i.e. that there be no quantitative distance between segments. Should there be distance, the least amount of this is better. The latter constraint specifies that each segment should provide the maximum number of qualitative differences, thus aiding in perception of different sounds. Of course, the actual number of rankings for either constraint varies from language to language. From this consideration, and given that categorization is a relative phenomenon established by each system, the following universal ranking hypotheses are advanced:

*MINDIST 2 (x) >> *MINDIST 1 (x)

> "It is less bad that a given segment be quantitatively distant from all others by a factor of 1 than by a factor of 2"
>
> MAINTAIN 1 (x) >> MAINTAIN 2 (x)
>
> "It is better that a segment be qualitatively contrastive from all others by two factors than by one factor"

Table 2.2 presents the interaction of constraints in a theoretical system containing the segments X, Y, and Z. In this example system, $X \times Y = 1$ quantitative distance, 1 qualitative contrast; $X \times Z = 2$ quantitative distances, 2 qualitative contrasts; and $Y \times Z = 0$ quantitative distances, 1 qualitative contrast.

{X, Y, Z}	MAINTAIN (1)	MAINTAIN (2)	*MINDIST (2)	*MINDIST (1)
X		*	*	**
Y		**		*
Z		*	*	*

Table 2.2. An Example of Perceptual Constraint Interaction

The system-specific interaction of segments in the theoretical system results in violation and satisfaction of perceptual constraints. The totality of these violations and satisfactions provides a description of the perceptual grammar of the theoretical system.

2.5. Functionalism and the rhotic question: the need for a relational phonology

Each of the drives associated with a functionalist understanding of human language approaches particular forms—in this case, phonological segments—from an inherently relational perspective. In describing the articulatory grammar of a theoretical inventory, satisfaction or violation of *HOLD constraints depends not on a system-external definition of length, but one that derives from the larger system. The perceptual drive and grammar depends exclusively on the interaction of system members: determination of quantitative distances and of qualitative contrasts cannot be accomplished without reference to this larger structure. In essence, functionalism and the constraint-based grammars I present here establish a relational approach to segments and inventories, where form is borne from an organic system whose constituents are at the same time distinct from and similar to each other.

Returning to the subject of rhotics, the inclusion of these segments in a given linguistic system raises two questions. If phonological inventories can be described using functionalist principles and constraint grammars, how do different rhotics fit into these systems? Also, given a functionally grounded, relational understanding of these and other segments, how can rhotic variation be described, explained, and predicted using a functionalist understanding of human speech? The first question looks to relational patterns for the definition of rhotics as a distinct type of phonological segment; the second looks to the particular behavior of rhotics.

A functionally inspired, relational approach to the question of rhotic phonology seeks to understand the articulatory and perceptual properties of each rhotic within an organic whole. In contrast to many of the works cited here, a relational phonology of rhotics—or of any other segment—cannot focus solely on articulation/production or on acoustics/perception, but must take both factors into account. Also in contrast to many of the studies cited in 2.3 and 2.4, it is not sufficient, within the context of a relational phonology, to consider only one articulatory or acoustic property of rhotics in seeking to describe or explain their apparent classhood. Before any discussion of phonological systems can be undertaken within such a perspective, systems and their respective members must be described phonetically, i.e. grounded in observations of tactile, gestural and acoustic reality. From phonetic evidence, relational definitions of segments may be advanced—both articulatory and perceptual—and instances of variation or particular phonological behavior analyzed. Relational phonology, as a working theory evolved in this dissertation, is therefore the natural extension of functional phonology (seen in Boersma) and of such phonetically inspired works as Flemming, Padgett, and Schwartz et al.

CHAPTER THREE. DESCRIBING CONTINUANT CONSONANT INVENTORIES

3.1. Introduction

This chapter focuses on the description of human speakers' continuant consonant inventories, looking at rhotics as well as at the larger system of gesturally and acoustically related sounds. I present data from a study involving speakers of three languages—American English, European French and two northern forms of Dutch—and describe continuant consonant sound segments in an organic (that is, natural language) context. My analyses and their articulation in this chapter make no pretension of phonetic exhaustiveness; I wish only to provide a proper phonetic basis for the phonological description of the systems and segments in question. Subsequent chapters integrate the resulting data into a theoretical approach to speech segments, a hypothesis about the interactive nature of sound systems that I term relational integrity.

The choice of languages for this study and the specific, dialectal forms targeted are partially motivated by pragmatic concerns, such as the availability of native speakers, but more generally by the rhotic and continuant consonant inventories of each language. The subjects used for this survey provide a sampling of rhotics in different languages that are genetically related, but whose continuant consonant inventories differ in complexity and content. The most simple inventory—in terms of sheer number of continuant consonants—is seen in

European French, with five members {l, s/z, ʒ, f/v, and ʁ}.[10] Standard American, the form of English used by most media and public officials in the United States, consists of six members {l, s/z, f/v, ʃ, ɹ, ð/θ}, and presents two distinctions: the lack of voicing opposition for the post-alveolar [ʃ] and the set of interdental or dental fricatives [ð, θ]. Dutch (or Netherlandic) presents a different six-member set, consisting of {l, f/v, s/z, ʃ, x, R}.[11]

3.2. Goals and Methodology

The goals of the experiment were to elicit instantiations of continuant consonants within the context of native speaker utterances, and to provide for the direct acoustic and indirect gestural description of each member of this sound inventory. A secondary goal was the generation of rhotic data, again within the context of utterances native to the target languages, providing phonologically oriented, systemically defined gestural and acoustic profiles of each r-like segment.

[10] The reader will notice that, throughout this dissertation, I follow a convention of French linguistic studies in using the IPA symbol [ʁ] in reference to the French rhotic. This symbol should be read as meaning, "a segment of voiced or voiceless quality, articulated in an approximant or fricative manner at a velar, uvulo-velar or uvular place of articulation."

[11] Given the variability of the Netherlandic rhotic, I use the symbol R when making generic reference to this segment. IPA symbols {ʀ, ʁ, ɹ, ɾ} are used in discussion of specific rhotics and their phonological variation.

3.2.1. DATA

The data for this descriptive experiment consisted of words from each of the three subject languages, containing members of the continuant consonant inventory in three distinct phonotactic environments: word-initial, intervocalic, and word-final. A number of considerations were taken into account in the development of stimuli. Firstly, glides or semi-vowels were not included in the stimulus sets. These segments behave phonologically like consonants, but their phonetic resemblance to vowels suggests that they should be considered as a separate sub-inventory. Secondly, the stimuli consisted of voiced segments whenever possible and the reader will notice a bias towards the presentation of these in spectral figures and charts. The voiced-voiceless distinction is, if present for a particular segment in the subject languages, a binary characteristic, whether considered in terms of the phonology or phonetics (note that I ignore questions of VOT). Language-specific distribution of voiced and voiceless segments resulted in asymmetries in the stimuli. The Dutch [ʃ] is seen only in loan words or as a result of morpheme concatenation in intervocalic environments and is not licensed in word final environments, save for very recent borrowings, most of which come from English (e.g. *hasch*). These were excluded from the stimulus set. Likewise, the Dutch [v] is obligatorily voiceless in word final environments, as seen in the opposition *dief* – *dieven*. Such asymmetries are also noted in other positions in modern Dutch spoken in the Netherlands, where the voiced member of fricative pairs is increasingly in decline, particularly in Amsterdam and surrounding areas. Another asymmetry is noted in American English. With very few exceptions, word-initial [z] occurs in play or nonsense words and in words of non-English origin: word-initial [s] was therefore used in all stimuli. Additionally, American English does not, in opposition to French, have a voiced counterpart to the post-alveolar, i.e. [ʒ], in all positions.

Words comprising the stimulus sets were chosen because they contained one of the target continuant consonants. For each continuant consonant, three words were selected, consisting of the target continuant adjacent to a high front, high back and low vowel, respectively. Every effort was made to use words of relatively high frequency, avoiding so-called exotic vocabulary where the subject might focus more on the stimuli and produce an overly articulated sample. Dutch stimuli consisted of the set {ʀ, l, z, v, x, ʃ} in 51 words. French stimuli consisted of the set {ʁ, l, z, ʒ, v} in 45 words. English stimuli consisted of the set {ɹ, s/z, l, ʃ, v, ð/θ} in 54 words. Several tokens in each language contained more than one continuant consonant, each of which was analyzed separately. The total number of continuant consonant measurements per speaker was 66 for English, 61 for Dutch, and 54 for French, resulting in a total of 362 continuant consonant measurements in the experiment. The list of stimuli used for the experiment is provided in Appendix A.

Two native speakers of each language were recorded in a quiet, semi-soundproof room. Speakers were asked to produce the stimuli, which had been randomized on a series of flashcards. Recording of speaker output was accomplished with an IBM compatible computer, using a high-quality digital input microphone. A sampling rate of 22050 Hz at a 4mB recording buffer was specified on Praat (version 3.8.24) speech analysis and synthesis software. Recordings were saved as digital sound files (*.wav) for future analysis and inclusion as speech samples in this dissertation. The original waveforms were modified to reduce background noise (from fluorescent lighting and computer fan hum) by filtering the waveform of this noise, then augmenting higher frequencies (treble boost) and finally by amplifying the original signal. Subsequent analysis of speech tokens was accomplished using Wavesurfer (version 1.0.4) software.

A distinction must be drawn in the calculation of formants and formant averages between fricatives and vowel-like segments. The former display qualitatively dispersed noise at frequencies as high as 8000 Hz; the later display relatively distinct, clear formants (generally below 4000 Hz) and are analyzed in much the same way as vowels. The reader will note that formant values for fricatives are more approximate—as they result from the averaging of spectra—and all formant measurements were rounded to the nearest 100 Hz. When determining formant values, only the most intense concentrations of energy in the overall fricative spectrum are considered and less intense frequencies are ignored. Unlike vowels, this does not look from the bottom of the spectrum upwards, but across the entirety of spectral energy. This approach to fricative formant calculation corresponds to the methodology used in phonetic literature.

Analysis of articulatory gestures is heavily dependent on existing literature and relies on a more fluid interpretation of the data. Whereas acoustic output can be measured with degree of precision, thus translated into a perceptual mapping of each segment, evidence for gestural output is less quantitatively precise (one gesture is rarely replicated exactly) and indirect (measurements rely on observation of secondary characteristics). Output acoustic patterns are used to provide an articulatory account for the segments in question by matching similar patterns to known gestural movements.

3.2.2. SUBJECTS

Subjects for this experience were adult (30-43 years) female native speakers of one of the target languages. The author solicited subject participation, as all potential participants were colleagues and/or personal acquaintances at the University of Texas or in the surrounding community. English speakers are all from the central or midwestern region of the United States and self-identified speakers of Standard American English.[12] The two Dutch speakers are both Netherlands' citizens and bilingual speakers of English. Subject N3 is from the Amsterdam region and identifies herself as an Amsterdammer, although she reports that her normal speech is less heavily influenced by regionalisms than is typical for a person from this city. Subject N4 is from Northern Brabant and identifies her speech as typical of this region.[13] Both French speakers are from Metropolitan France and speak a standard form of European French, with little discernable regional accents. Subject F5 identifies herself as Parisian; subject F6 is from Lyon.

---

[12] Each English subject was asked whether she believed her manner of speaking and accent was similar to media personalities and relatively free of regional accents.

[13] References to the speech of subject N4 as being "Brabant Dutch" should not be confused with "Brabants." The former term is used to describe the local form of standard or shared Netherlandic spoken in Dutch Brabant; the latter term refers to this region's dialect or dialect-group.

3.3. Results

The results of this experiment are presented in seven sections. The first four consist of continuant consonants common to each of the subject languages: [s, z], [ʃ, ʒ], [f, v], and [l]. The next two sections describe segments found in only one language: [ð, θ] of English and [x] of Dutch. A final section examines the tentatively grouped rhotics. In each, the qualitative and quantitative acoustic profile of segments is discussed and equated to gestural patterns. Example spectra are given as a visual representation of the acoustic profile of segments. For fricatives, I have chosen to display example spectra—both FFT and LPC—for each segment, taking into account voiced and voiceless variants, as well as averaged spectra for each speaker and for each target environment. Because laterals and rhotics present much clearer, vowel-like formants—even in the case of fricative articulations—I have presented only FFT and LPC spectra of one token per speaker and per target environment. Appendix B provides recordings, waveforms, spectrograms, and spectra (FFT and LPC) for all tokens: these may be accessed with any Windows based web browser that is Java enabled.

3.3.1. THE ALVEOLAR SIBILANT [s, z]

Each of the subject languages contains the sibilant fricative [s, z]. In either voiced or voiceless form, this segment is present in all three phonotactic environments studied. The most general—and most prominent—acoustic characteristic of [s, z] is the widely dispersed energy seen at and above approximately 4000 Hz—most often near 5000 Hz—and extending to approximately 8500 Hz. This energy is, in each of the subject languages, most intense at upper levels, approximately 5500-7000 Hz. Some energy is noted below 4000 Hz, especially in intervocalic instantiations: this is attributed to dampened source energy and is not to the fricative gesture. Given the wide dispersion of energy, as well as the dampening characteristic of all fricatives, distinction of spectral peaks is an ambiguous, rather than exact task. It is nevertheless possible to provide approximations of areas of energy intensity, which I term consonant formants, abbreviated as F1, F2, and F3 for the three most prominent regions of acoustic energy.[14]

For all subjects, three regions of acoustic prominence were observed. The first of these occurs between 4000 and 5500 Hz and presents one of the more intense regions of energy. F2 is seen at roughly 6000 Hz and is also quite intense, often more so than F1. The intensity of F3 varies considerably, as discussed below, and is typically seen between 7200 and 8000 Hz. Figures 3.1.a and b show sample spectra of [z] and [s], respectively.

[14] Shadle, Badin and Moulinier (1991) also use the term formant to refer to the spectral peaks of fricative consonants. Stevens (1999) makes use of this term in some discussion of fricatives, as do Kent and Read (1992).

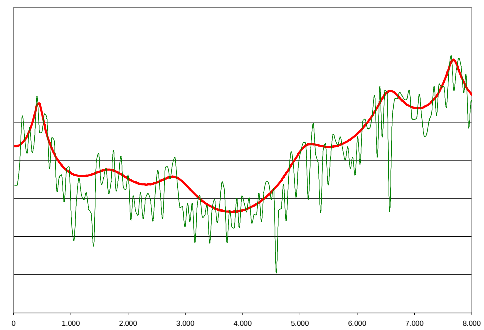


Figure 3.1.a. Example FFT (green) and LPC (red) spectra of [z]: F5 token *zaïre* [za.iʁ].

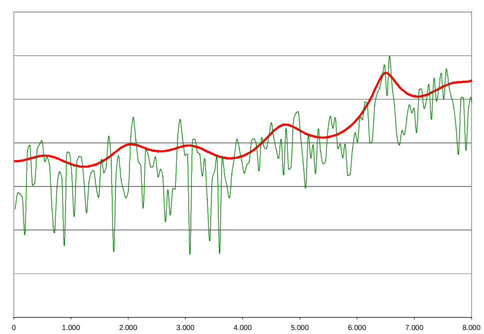


Figure 3.1.b. Example FFT (green) and LPC (red) spectra of [s]: E2 token *sat* [sæt].

3.3.1.1. English

Both English speakers produced [s] and [z] tokens with widely distributed formants. Qualitatively, the sibilant fricatives of both subjects E1 and E2 showed were quite similar: both consisted of widely dispersed energy from approximately 5000 to 8000 Hz. Subject E1 tokens display considerably more dampening in intervocalic instantiations, a trend not as noticed in the tokens produced by E2.

Speaker E1's tokens show very little phonotactic variability, save for the slight raising of F2 in intervocalic instantiations. Quantitative measurement for this speaker provides for average formant values of 4925 Hz, 6183 Hz, and 7883 Hz in word-initial tokens; 5000 Hz, 6317 Hz, and 7800 Hz in intervocalic tokens; and 5013 Hz, 6100 Hz, and 7850 Hz in word-final tokens. Sample spectra for this speaker are given in Figure 3.2, below.

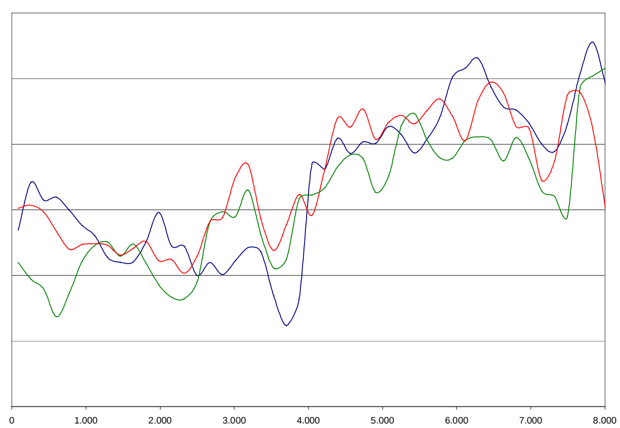


Figure 3.2. Averaged spectra of sibilant fricative [z] as produced by speaker E1: *sat* [sæt] (green); *pausing* [pa:ziŋ] (blue); *pause* [pa:z] (red).

Subject E2's tokens are quantitatively notable for the same lack of variation as seen in E1: here the most important exception is F3 in word-initial instantiations. In these environments, F3 is somewhat greater, although this difference is not tremendous in terms of the relative stability of higher formants. E2 produced tokens whose average formants measured 4792 Hz, 6343 Hz, and

7683 Hz in word-initial samples; 4650 Hz, 6300 Hz, and 7917 Hz in intervocalic samples; and 4688 Hz, 6488 Hz, and 7875 Hz in word-final samples. Figure 3.3 provides contrastive profile spectra for this speaker.

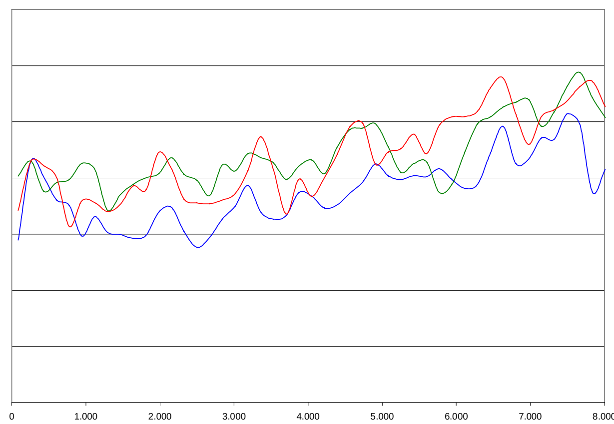


Figure 3.3. Averaged spectra of sibilant fricative [z, s], speaker E2: *sat* [sæt] (green); *pausing* [pa:ziŋ] (blue); *pause* [pa:z] (red).

3.3.1.2. Dutch

Both Dutch speakers produced [z] and [s] with little variance. Dutch presents a number of particularities, vis-à-vis the other languages of this study. Compared to the English speakers, a greater degree of overall fricative intensity is noted, although this energy is—as with all other speakers—extremely distributed from 4000 Hz to 8000 Hz. Voicing characteristics varied between the two Dutch speakers. For N3, all word-initial instantiations were voiced; for N4, these were consistently voiceless. In intervocalic environments, voicing varied for both N3 and N4: both speakers showed a marked tendency toward devoicing of the sibilant, although it is difficult to determine to what extent residual voicing is the result of adjacent vowels. As mentioned above, the sibilant is always voiceless in word final environments: no exceptions were noted in speaker data.[15]

Also noted in the output of both speakers was a great degree of coarticulatory variance, especially in those tokens that were produced prior to or immediately following a back vowel, [u] and [o]. For these tokens, the spectral profile of the consonant was significantly tilted towards lower frequencies. By contrast, the profile of tokens adjacent to higher, front vowels [i] and [e] shows a bias to higher frequencies.

Speaker N3 produced some of the most intense tokens of any included in this study, in terms of relative prominence of formant peaks (measured in dB). Especially in word-initial environments, dense regions of acoustic intensity are noted from approximately 4000 Hz to 8000 Hz and beyond. Intensity is relatively less in intervocalic and word-final instantiations. Formant peaks for this speaker were measured at average points of 4720 Hz, 6280 Hz, and 7610 Hz in word-initial environments; 4500 Hz, 6150 Hz, and 7567 Hz in intervocalic environments; and 4500 Hz, 6188 Hz, and 7625 Hz in word-final environments. Figure 3.4 provides a visual image of average spectra of [z] for this speaker.

[15] One particularity noted in the data is that of speaker N4's articulation of *boos*. This segment more closely resembled that of the post-alveolar [ʃ]. This variation may be due to a regional pronunciation and this token was not included in the analyzed data set.

Figure 3.4. Averaged spectra of sibilant fricative [z, s], speaker N3: *zaad* [za:d] (green); *azen* [ɑzən] (blue); *daas* [da:s] (red).

Little difference was noted in tokens of speaker N4, save for a relative lowering of all formants with regard to those of N3 and the devoicing, discussed above. Average formant values for this speaker were calculated at 4530 Hz, 6210 Hz, and 7660 Hz in word-initial environments; at 4383 Hz, 6083 Hz, and 7683 Hz in intervocalic environments; and at 4933 Hz, 6400 Hz, and 7817 Hz in word-final environments. Sample spectra for this speaker are given in Figure 3.5.

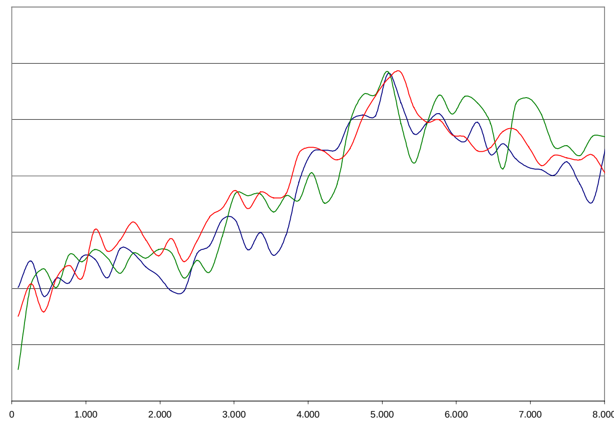


Figure 3.5. Averaged spectra of sibilant fricative [s, z], speaker N4: *zaad* [za:d] (green); *azen* [ɑzən] (blue); *daas* [da:s] (red).

3.3.1.3. French

French presents the case of a language having the voiced sibilant in all environments, in contrast to English and Dutch. The data was not, however, without irregularities: speaker F6 systematically devoiced the sibilant fricative in word-final positions. This variation is not seen in speaker F5 and represents a particularity for which I have no explanation. Differences were also noted in the relative intensity of each speaker's output. Speaker F5 regularly produced more robust and higher-frequency tokens—in all environments and for all stimuli—than did F6. Lacking any explanation for such variability in phonetic or phonological literature, I attribute this variation to individual speaker style.

A difference noted throughout this study, which is also mentioned in subsequent contexts, is the nature of French consonant-to-vowel and vowel-to-consonant coarticulatory effects. The acoustic profile of French consonants appears to be highly influenced by adjacent—especially tautosyllabic—vowels. Coarticulatory effects are seen in both French speakers, although this is more evident in the output of F5 than of F6: again, this must be attributed to speaker style, perhaps also to more careful articulation on the part of speaker F6. Regardless of the extent of vowel influence on consonants, a number of interesting variations are noted in the data. Adjacent to [i], all formants are relatively greater; correspondingly, those segments adjacent to [u] are lower. Medial measurements are noted in [a] adjacent environments.

Speaker F5 produced some of the highest frequency formants of any in this study, for all formants and in all environments. Average formant values for F5 were calculated as 4817 Hz, 6467 Hz, and 7950 Hz for word-initial tokens; as 5386 Hz, 6538 Hz, and 7838 Hz for intervocalic tokens; and as 5000 Hz, 6467 Hz, and 7733 Hz for word-final tokens. Figure 3.6 provides examples of averaged spectra of [z] for this speaker

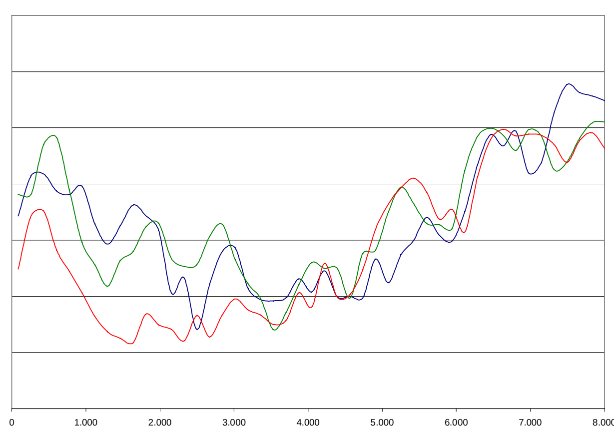


Figure 3.6. Averaged spectra of sibilant fricative [z], speaker F5: *zaïre* [za.iʁ] (green); *bazar* [bazaʁ] (blue); *base* [baz] (red).

Speaker F6 produced tokens whose average formant peaks were measured at 4850 Hz, 6217 Hz, and 7467 Hz for word-initial instantiations; at 5038 Hz, 6350 Hz, and 7488 Hz for intervocalic instantiations; and at 5050 Hz, 6183 Hz, and 7350 Hz for word-final instantiations. Figure 3.7 presents contrastive, averaged spectra for [z] as produced by F6.

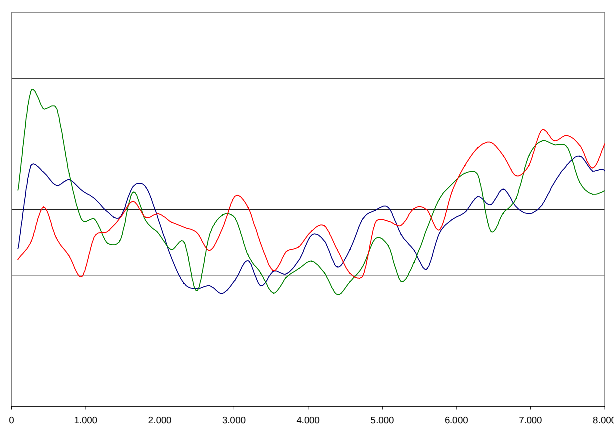


Figure 3.7. Averaged spectra of sibilant fricative [z], speaker F6: *zaïre* [za.iʁ] (green); *bazar* [bazaʁ] (blue); *base* [baz] (red).

3.3.1.4. Gestural account

The higher frequency fricative energy noted in the acoustic output of alveolar sibilants is the result of airflow through a constricted region in the central part of the oral cavity. Frication occurs at the aperture created between the tongue and the alveolar ridge, sometimes involving the teeth or the post-alveolar region. Two gestural particularities distinguish [s, z] from other alveolar fricatives, most notably [ð, θ]. These are jaw height and tongue contour.

Keating (1983) discusses the relative height of the jaw for both stop and continuant consonants in American English. She notes that for [s] in particular, the jaw is nearly or entirely clenched during the consonant gesture, as much as or more than for glides, nasals, and stops. Unfortunately, her data does not include other fricatives. A secondary consideration is the contour of the tongue, which varies quite naturally from speaker to speaker. In a study of 20 American English speakers, Dart (1991) notes that 52.5% of instantiations were made using the tongue blade (lamina) and 42.5% using the tongue tip. Ladefoged (1957) presents an illustration of his own articulation of [s] in *saw*: this is given below in Figure 3.7. Here, the tongue is in a grooved or "U" shape, with lateral contact maintained along the sides of the palate. The gestural target is the alveolar region, directly posterior to the front incisors. In a later study, Ladefoged and Wu confirmed the relationship of gesture to acoustic signal in the study of Pekingese Mandarin: although precise tongue contour varies widely between speakers, output acoustics are consistent (1984: 269-70).

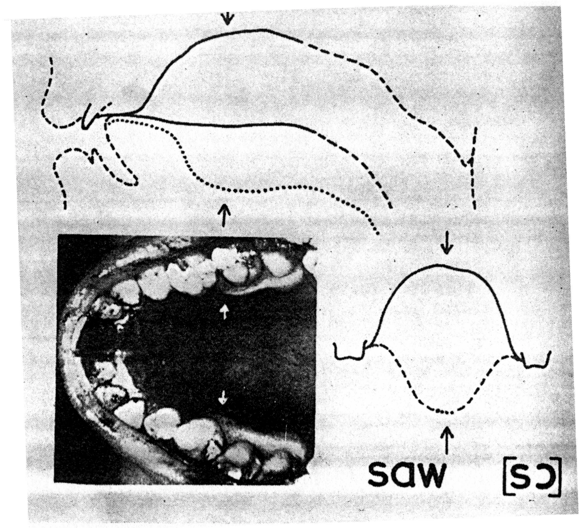


Figure 3.8. Gestural picture of [s] (Ladefoged 1957: 770).

Other studies confirm the tongue position and contour and the gestural target observances presented for English. Wängler (1976) provides x-ray photographs from German similar to the tracings of Ladefoged and Ladefoged and Wu: his sole subject produced [s] with a grooved tongue shape and the tongue blade targeting the alveolar ridges. Pétursson (1971) shows another example from Icelandic; here, [s] is produced with the tongue tip and in closer contact with the upper front incisors. Bladon and Nolan (1977) offer a different means for quantifying gestures. In their study, which focuses on alveolars in English, they provide distinction between three sets of consonants, [s, z], [t, d, n], and [l] according to two parameters: blade height and tip height. Each of these varies along an indexed continuum of 1 to 3. Figure 3.9 shows the stylization of these parameters: gestured indexed as 3-1 (blade and tip, respectively) are [s, z]; those 3-2 and 3-3 are [t, d, n]; and those 1-3 and 2-3 are

/l/ (1977: 189, 192). Their analysis suggests that the tongue blade is a primary articulator, though the tongue tip may be implied, albeit to a lesser extent, in some instantiations and by some speakers.

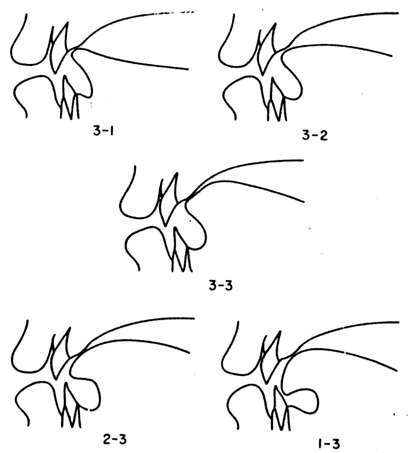


Figure 3.9. Tripartite gestural description of alveolars (Bladon and Nolan 1984)

Given the descriptions of sibilants in a variety of languages, it is possible to paint a much broader gestural picture. This is not intended to account for every gesture noted in the speakers, but to provide a means of contrasting the significant movements involved in segmental production. The primary articulator involved in [s, z] is the tongue front (the tongue blade or tip, an unimportant distinction in this instance). The target of articulator movement is the alveolar ridge, where the resulting constriction produces frication and resulting noise. The shape of the tongue is grooved (i.e. contracted laterally) and there is no periodicity, as seen in trills or flaps, as the constriction is maintained for the duration of the gesture

3.3.2. THE POST-ALVEOLAR (PALATAL) SIBILANTS [ʃ, ʒ]

Each of the subject languages has in its inventory a post-alveolar or palatal sibilant. In two of the languages—English and Dutch—only the voiceless [ʃ] is present, in the latter only in word-initial and intervocalic environments as a result of morpheme concatenation. The voiced [ʒ] is seen in all environments in French.

Many qualitative similarities between alveolar [s, z] and post-alveolar [ʃ, ʒ] are noted in the data. The acoustic output of both segments is, in all languages and in all contexts, greatly dispersed and distributed over a large part of the spectrum. In contrast to the alveolar sibilants, post-alveolars show more concentrated energy in the region of 3000-4500 Hz. The first region of spectral prominence is in most cases seen at approximately 3000 Hz, with the second, more robust formant at approximately 4000-4500 Hz. A third spectral peak is found in the region of 5500-6500 Hz.

Measurement of sound quantity—i.e. the measurement of spectral peaks or formants—presented a challenge in the case of post-alveolars. Because of the concentrated, yet dispersed nature of the output energy, determination of consonant formants was even more approximant than in the case of alveolar sibilants. This should not be taken as a shortcoming in the data or measurement, however: the dispersed, distributed nature of post-alveolars is clearly one of their most salient perceptual characteristics and it may well be that this variability is what leads listeners to correctly map the output to a lexical specification. Figures 3.10.a and b show example spectra of the post-alveolar sibilants.

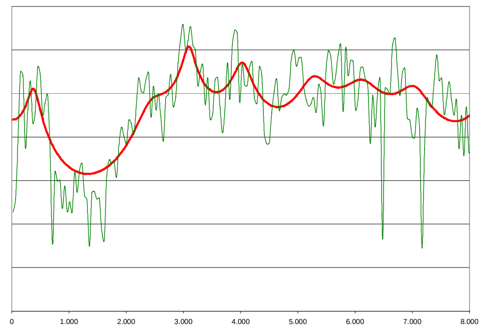


Figure 3.10.a. Example FFT (green) and LPC (red) spectra of [ʒ]: F5 token *j'ai* [ʒe].

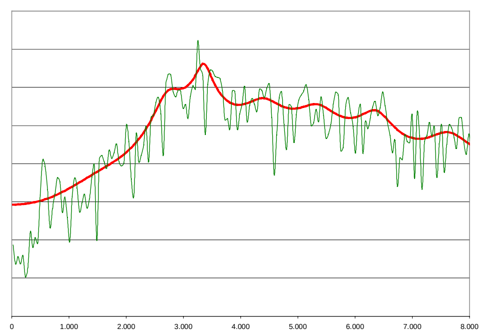


Figure 3.10.b. Example FFT (green) and LPC (red) spectra of [ʃ]: N3 token *sjaat* [ʃa:t].

3.3.2.1. English

Both English speakers produced [ʃ] tokens with little phonotactic variation save for relatively shorter articulations in intervocalic instantiations. Formant peaks were measured at somewhat greater frequencies in these environments. In nearly all environments, speaker E2 produced tokens with greater formant values than did E1. As opposed to [s, z], where higher formants were more robust in all environments, lower F1 and, especially, F2 are characteristically the most salient spectral peaks in [ʃ] instantiations. Widely distributed energy above 4500 Hz is noted in all tokens. In most cases, the third region (at ±6000 Hz) of spectral prominence is only marginally more intense than these surrounding regions of prominence.

Tokens produced by speaker E1 resulted in average formant values of 2883 Hz, 4050 Hz, and 6067 Hz in word-initial environments; of 3117 Hz, 4167 Hz, and 6017 Hz in intervocalic environments; and of 2933 Hz, 3900 Hz, and 5850 Hz in word-final environments. Figure 3.11 presents averaged spectra of [ʃ] in each of the three target environments.

Speaker E2 produced tokens whose average formant values were 3083 Hz, 4050 Hz, and 6067 Hz for word-initial tokens; 3083 Hz, 4400 Hz, and 6500 Hz for intervocalic tokens; and 3033 Hz, 4133 Hz, and 6250 Hz for word-final tokens. Figure 3.12 presents example spectra of three instantiation of [ʃ] by speaker E2.

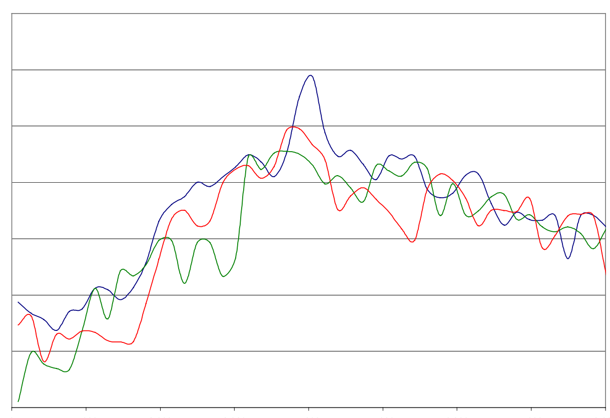


Figure 3.11. Averaged spectra of post-alveolar fricative [ʃ], speaker E1: shall [ʃæl] (green); caution [kaʃən] (blue); gosh [gɑʃ] (red).

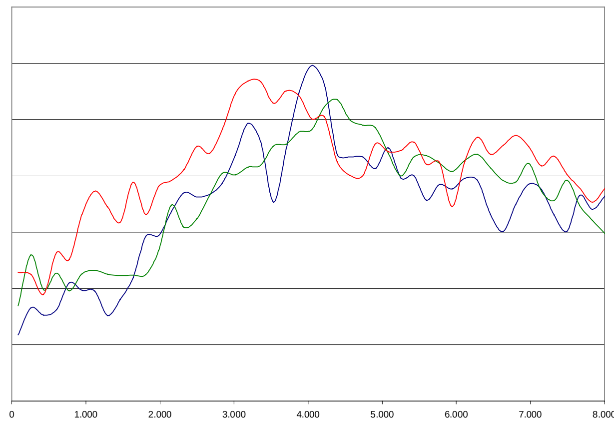
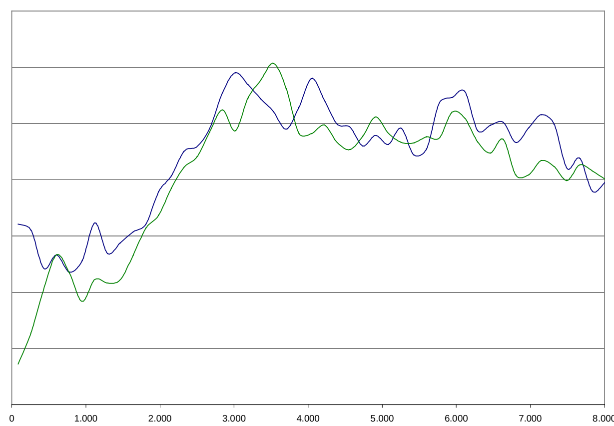


Figure 3.12. Averaged spectra of post-alveolar fricative [ʃ], speaker E1: shall [ʃæl] (green); caution [kaʃən] (blue); gosh [gɑʃ] (red).

3.3.2.2. Dutch

The Dutch post-alveolar sibilant [ʃ] is, for both speakers, similar to that of English, in that nearly all of the output energy is concentrated above 2000 Hz, although there is greater and more intense energy above 5000 Hz than noted for either English speaker. Additionally, there is much greater variation in [ʃ] production in Dutch than in English, both inter- and intra-speaker. It is not clear to what extent the non-nativeness of [ʃ] in Dutch can account for the speaker- and token-specific variation of the post-alveolar sibilant.

As noted for [z, s] production, speaker N3 produced [ʃ] with characteristically greater intensity than did either her compatriot N4 or either of the English speakers. N3 tokens display widely distributed energy above 2000 Hz, extending well into the region of 7000-8000 Hz. The most intense energy is seen between approximately 2500 Hz and 4000 Hz. Production of [ʃ] by this speaker resulted in average formants of 2817 Hz, 3800 Hz, and 6117 Hz in word-initial environments and of 2883 Hz, 3967 Hz, and 6100 Hz in intervocalic environments. Figure 3.13 presents averaged spectra from two such tokens of speaker N3.



3.13. Averaged spectra of post-alveolar fricative [ʃ], speaker N3: *sjaat* [ʃa:t] (green); *meisje* [mɛʲʃɘ] (blue).

In most environments, speaker N4 produced tokens of relatively less intensity and at relatively lower frequencies than did speaker N3. Also noted in this speaker was a lack of relative intensity above ±6000 Hz, making her tokens more similar to those of E1 and E2. For this speaker, average formant values were calculated at 2617 Hz, 3950 Hz, and 6000 Hz for word-initial tokens, and at 2850 Hz, 3833 Hz, and 5667 Hz for intervocalic tokens. Represented in Figure 3.14 are contrastive examples of [ʃ] production for speaker N4.

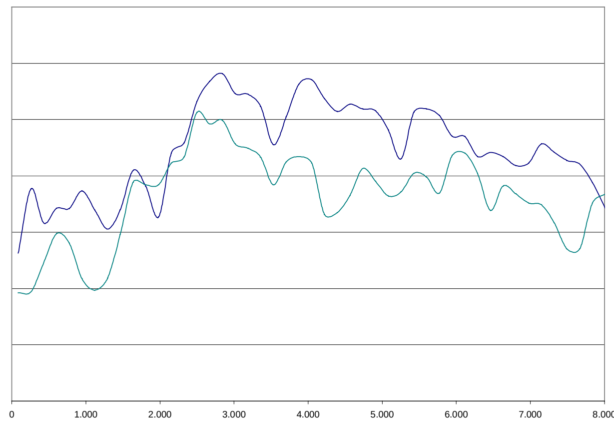


Figure 3.14. Averaged spectra of post-alveolar fricative [ʃ], speaker N4: *sjaat* [ʃa:t] (green); *meisje* [mɛʲʃɘ] (blue).

3.3.2.3. French

The French [ʒ] is, despite differences in voicing, very similar to both English and Dutch. For both speakers, the quality of output energy consisted of dispersed energy distributed over a large area extending from approximately 2000 Hz to 6500 Hz. Less robust energy extended to and beyond 6500 Hz for both speakers. As noted in 3.3.1.3 in discussion of [z], vowel-to-consonant coarticulatory interference is observed for both speakers, although the extent of this influence is significantly less in the case of [ʒ].

As with her production of [z], speaker F5 produced tokens whose intensity—namely in the upper formant regions—was somewhat greater than speaker F6. Save for this one qualitative feature, little difference was noted between French speakers. For F5, average formant measurements were calculated at 2950 Hz, 3967 Hz, and 6200 Hz for the consonant in word-initial environments; at 2825 Hz, 3875 Hz, and 6375 Hz in intervocalic environments; and at 2933 Hz, 4183 Hz, and 6183 in word-final environments. Figure 3.15 shows averaged spectra of [ʒ] for this speaker.

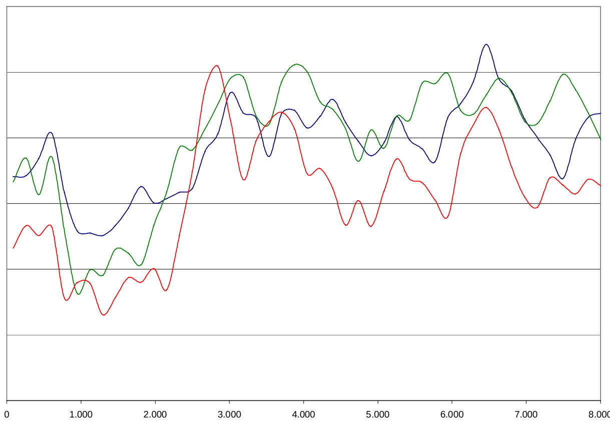


Figure 3.15. Averaged spectra of post-alveolar fricative [ʒ], speaker F5: *j'ai* [ʒe] (green); *gagea* [gaʒa] (blue); *cage* [kaʒ] (red).

Speaker F6 produced tokens of a relatively lessened intensity. Although acoustic energy was noted above 6000 Hz for this speaker, this was not as intense as that seen for speaker F5. Average formant measurements for F5 are given as 2883 Hz, 4100 Hz, and 6150 Hz for word-initial tokens; as 2725 Hz, 4100 Hz, and 6225 Hz for intervocalic tokens; and as 2917 Hz, 4133 Hz, and 6217 Hz for word-final tokens. Three examples of F6's [ʒ] are seen in the averaged spectra of Figure 3.16.

Figure 3.16. Averaged spectra of post-alveolar fricative [ʒ], speaker F5: *j'ai* [ʒe] (green); *gagea* [gaʒa] (blue); *cage* [kaʒ] (red).

3.3.2.4. Gestural account

The post-alveolar sibilants are gesturally distinct from their alveolar counterparts. For these, a reduced aperture is achieved in the oral cavity at a point of constriction posterior to the alveolar ridges. The turbulent airflow passes through this aperture at the central part of the oral cavity, similarly to [s, z]. This gesture involves the tongue front—either blade or tip—and the large dome of the palate. Ladefoged (1957) suggests that not only are [ʃ, ʒ] distinguished by the gestural target but also by the shape of the tongue. Whereas alveolars are produced with a grooved tongue, constricted in a u-like shape, post-alveolars are articulated with a relatively flat tongue. Figure 3.17 shows these tracings, based on his x-ray study. Photographic and x-ray data from Wängler's (1976) subject confirms this positional distinction: his subject shows much broader tongue contact in the palatal ridges and a curious rounding of the lips that may be particular to German. Straka (1979) contrasts the French voiced [ʒ] with the semi-vowel [j], showing that the critical difference between the two articulations is the shape of the tongue. The former is produced with a relatively flat tongue, whereas the latter implies a significant contraction of the tongue body (148).

Figure 3.17. Gestural representation of [ʃ/ʒ]: the vertical arrow indicates the point of constricted airflow (Ladefoged 1996: 149).

Post-alveolar sibilants can be characterized as being articulated with the tongue front (in this case, the blade). The tongue is in a flat, or non-contracted shape and the target of distal movement is the post-alveolar or palatal region. This is accomplished in a non-periodic manner, as the constriction is maintained for the duration of articulation and does not result in an open-close movement of target or articulator.

3.3.3. THE LABIAL FRICATIVES [f, v]

The labial fricatives [v] and [f] are the only non-sibilant, fricative segments shared by all three subject languages. English and French maintain voiced-voiceless distinctions in all environments, whereas Dutch prohibits voiced segments in word-final position.

The labial fricatives are characterized by relatively flat spectra, as compared to [s, z] and [ʃ, ʒ]. As seen in 3.3.5, this flat spectral profile is shared with the dental fricatives [ð, θ] of English. The flat, highly dispersed spectral profile presents a number of tightly grouped spectral peaks whose relative intensity is not great with regard to the overall spectrum. For this reason, no formant measurements are given for [f, v]. Rather, I have provided data for the region of energy and for the portion of the spectrum that shows the more intense energy. In most cases, fricative energy was observed from 1000 Hz to 8000 Hz, extending to nearly 10,000 Hz in many cases. While it would be exaggerated to say that the peaks at any frequency region stands out by virtue of regular intensity (such as ±3000 Hz for [ʒ] and ±5000 Hz for [z]), a relative intensity is observed at 7000-8000 Hz in most labial fricative instantiations. Figure 3.18.a and b provide sample spectra of labial fricatives.

The quality of energy produced by labial frication consists of extremely dispersed, higher-level energy. In intervocalic positions, especially, vowel-to-consonant transitions are observed, as are consonant formants of relatively greater intensity than in word-initial instantiations. I have included measurements of spectral peaks below ±8000 Hz; although significant energy is present above this level, the relational and perceptual significance of this is doubtful.

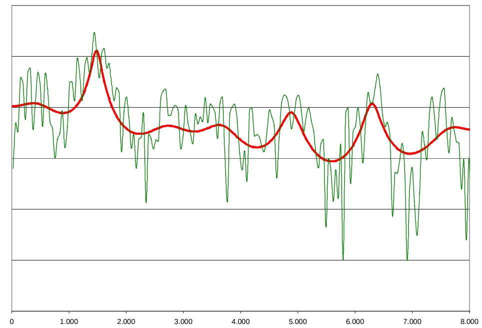


Figure 3.18.a. Example FFT (green) and LPC (red) spectra of [v]: F5 token *va* [va].

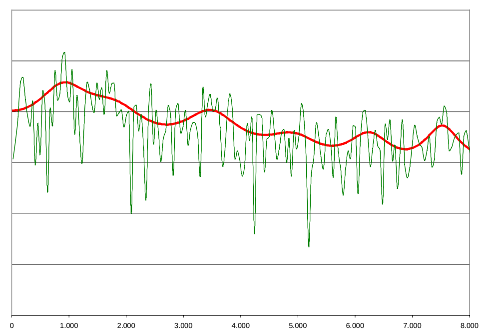


Figure 3.18.b. Example FFT (green) and LPC (red) spectra of [f]: N4 token *kaf* [kaf].

3.3.3.1. English

With a very few exceptions, English speakers produced [f, v] having very dispersed, dampened spectral peaks distributed over a range of approximately 5000-8000 Hz. Additional energy was observed above 8000 Hz for both speakers. Energy noted below 4500 Hz was most significant in voiced and intervocalic tokens and is the result of dampened resonance in the buccal cavities.

In these and all other cases, lower frequencies are not considered to be intrinsic to the acoustic profile of [f, v].

Tokens produced by speaker E1 varied little according to environment. The average range of relative intensity for E1 was calculated at 5188-7950 Hz in word-initial environments; 5117-8067 Hz in intervocalic environments; and 5333-8000 Hz in word-final environments. Figure 3.19 presents averaged spectra of [v] as produced by speaker E1.

Figure 3.19. Averaged spectrum of labial fricative [v], speaker E1: *vault* [vɑlt] (green); *coughing* [kɑfiŋ] (blue); *suave* [swa:v] (red).

Speaker E2 produced tokens similar to those of her homologue, the only remarkable distinction being E2's somewhat lower and less intense spectral peaks, especially in word-final tokens. Averaged intensity for E2 was calculated at 4938-7925 Hz in word-initial environments; at 4767-7917 Hz in intervocalic environments; and at 4467-7700 Hz in word-final environments. Figure 3.20 presents contrastive, averaged spectra of the labial fricative of E2.

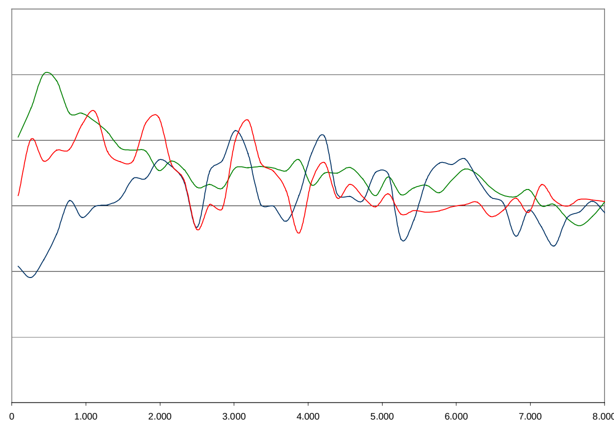


Figure 3.20. Averaged spectrum of labial fricative [v], speaker E2: *vault* [vɑlt] (green); *coughing* [kɑfiŋ] (blue); *suave* [swa:v] (red).

3.3.3.2. Dutch

Despite their relative quantitative similarity, Dutch tokens presented a number of qualitative distinctions, vis-à-vis those of English. Both Dutch speakers devoiced word-initial labial fricatives, as is characteristic for northern dialects of Netherlandic. Even more notable was the intensity of Dutch [f, v], in all environments greater than any produced by an English speaker. This was especially true of speaker N3. All spectral peaks were dampened and closely dispersed below 8000 Hz, with the most concentrated energy noted in the region of 4500-8000 Hz.

Speaker N3 showed tremendous variation in the intensity of speech tokens according to position. Word-initial tokens were significantly more intense than were any others; conversely, intervocalic and, especially, word-final tokens presented flatter, less robust spectral peaks. In all cases, N3 produced labial fricatives with closely bunched, relatively flat spectral profiles. Average measurements of fricative prominence were calculated at 4817-7800 Hz in word-initial environments; 4617-7750 Hz in intervocalic environments; and 4250-7363 Hz in word-final environments. Figure 3.21 gives averaged spectra of [ɣ̥] as produced by N3.

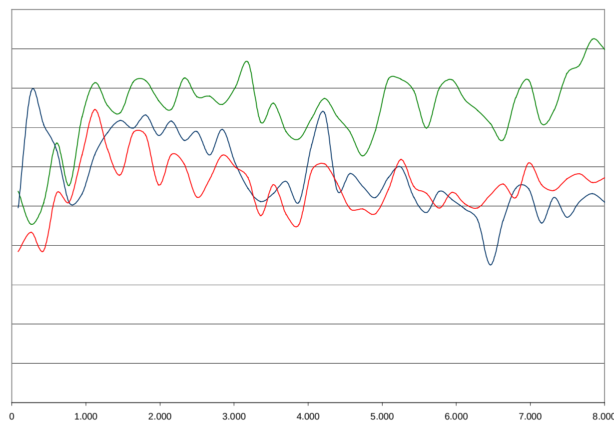


Figure 3.21. Averaged spectrum of labial fricative [ɣ̥], speaker N3: *vaag* [ɣ̥a:x] (green); *haven* [haɣ̥ən] (blue); *gaaf* [xa:f] (red).

As mentioned above, tokens produced by speaker N4 were relatively less intense than those of N3, although typically more intense than those of either English speaker. A notable feature of N4 tokens is the slightly greater intensity of energy at ±6000 Hz in word-initial tokens; this is not seen in other environments. Average values of spectral prominence for N3 were 6100-7950 Hz in word-initial tokens; 4483-7650 Hz in intervocalic tokens; and 4750-7423 Hz in word-final tokens. Contrastive examples of the spectral profile of N4 labial fricatives are given in Figure 3.22.

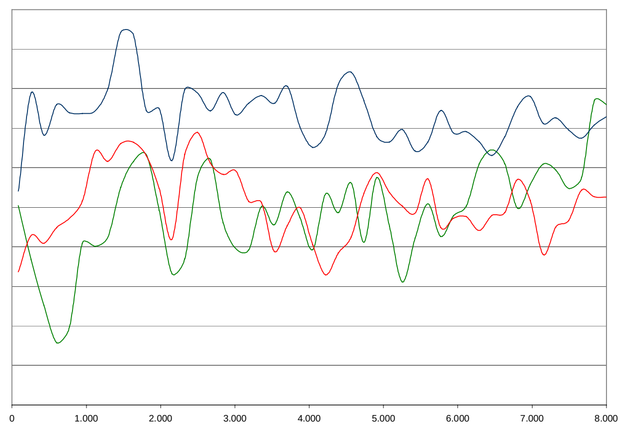


Figure 3.22. Averaged spectrum of labial fricative [v̥], speaker N4: *vaag* [v̥a:x] (green); *haven* [hav̥ən] (blue); *gaaf* [xa:f] (red).

3.3.3.3. French

All tokens from speakers F5 and F6 consisted of voiced variants. Much like corresponding tokens from English and Dutch, the formant observed in these samples consisted of highly dispersed fricative-like energy distributed over a flat spectrum extending above 8500 Hz. Qualitatively, the intensity of French [v] tokens resembles more closely those of English. Quantitative measurement revealed very little difference: tokens produced by F5 and F6 showed the same flat spectra as those of all other participants, with nominal peaks throughout the region of approximately 5000-8500 Hz. As is the case with other French consonant tokens, a large degree of vowel-to-consonant coarticulatory influence was noted. For both speakers, consonants adjacent to [i] presented characteristically higher energy; those adjacent to [u] resulting in correspondingly lowered spectra. Tokens adjacent to [a] yield intermediate profiles.

Speaker F5 produced relatively robust [v] tokens in all environments, with very little positional variation noted in any instantiation. Tokens of this speaker showed a significant amount of lower energy (-4000 Hz), most notably in intervocalic instantiations of the fricative. These are the result of resonance in the buccal cavity and are not considered to be significant acoustic components of the fricative itself, but an acoustic by product of the fricative gesture. For F5, spectral peaks are seen over an averaged spectral range of 5167-7300 Hz in word-initial tokens; 5350-7667 Hz in intervocalic tokens; and 5167-7617 Hz in word-final tokens. Figure 3.23 gives averaged spectra of [v] for this speaker.

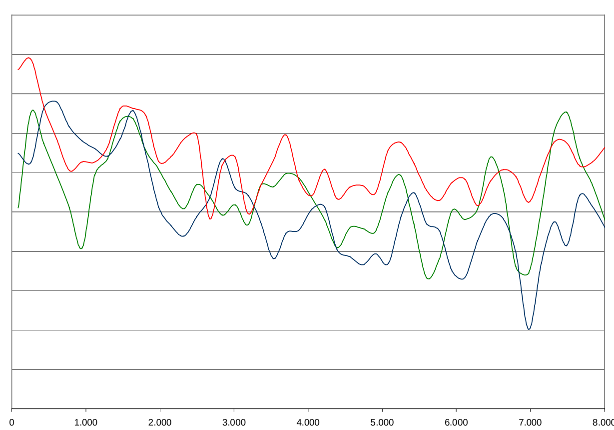


Figure 3.23. Averaged spectrum of labial fricative [v], speaker F5: *va* [va] (green); *avare* [avaʁ] (blue); *bave* [bav] (red).

While similar to those of F5, tokens produced by speaker F6 were somewhat less intense. F6 tokens presented average spectra with relative regions of intensity ranging from 5217 Hz to 7467 Hz in word-initial instantiations; from 5350 Hz to 7517 Hz in intervocalic instantiations; and from 5167 Hz to 7350 Hz in word-final instantiations. A contrastive example of [v] as produced by this speaker is given in Figure 3.24, below.

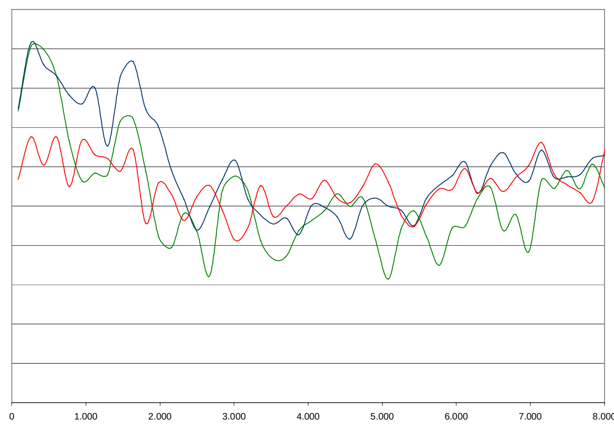


Figure 3.24. Averaged spectrum of labial fricative [v], speaker F6: *va* [va] (green); *avare* [avaʁ] (blue); *bave* [bav] (red).

3.3.3.4. Gestural account

The labial or labio-dental fricatives [f, v] are produced with the lower lip in close proximity and contact with the upper front incisors. The resulting disbursed fricative energy and flat spectral profile is the result of air movement between a tight but wide aperture and the internal resonance of the oral cavity. There is little literature suggesting important variation trends in the production of [f, v], save for mention of certain bilabial variants (Jones 1922, Gibson 1970, Nooteboom and Cohen 1995).

Labial fricatives are characterized as being produced by the lower lips, whose distal target is the upper front incisors. It is this near closure that produces the dispersed, higher frequency noise. The lack of prominence in [f, v] spectra is due to the dispersion of such noise, as there is no intervening secondary articulators and only secondary resonance in the buccal and glottal cavities. The shape of the tongue is not relevant to labial fricative production, nor is there any periodicity.

3.3.4. LATERAL [l]

All three of the subject languages allow the lateral [l] in each of the three target environments. In Dutch and English, word-final laterals are commonly velarized or "dark," transcribed as [ɫ]. No such variation is observed in French.

More than any other member of the continuous consonant inventory—save for certain rhotics, such as that of English—the laterals resemble vowels in both the quality and quantity of acoustic output. All [l] present clear, if dampened, vowel-like formants with very little of the dispersion noted in fricatives. For non-velarized or "clear" laterals, formants are typically discernable around ±400-1000 Hz (F1), ±1200-1800 Hz (F2), and ±2900-3500 Hz (F3). Velarized variants consisted of lowered F2 and, in the case of English tokens, the coordinated raising of F1. In all tokens, F3 was below 4000 Hz and very little energy was observed at higher frequencies.

Laterals are subject to vowel-to-vowel coarticulatory effects, especially in French, where F2 and F3 are influenced by adjacent vowels. The pattern of vowel-to-consonant interference is comparable to that seen in other buccal continuant consonants. This phenomenon is not significant in either English or Dutch.

3.3.4.1. English.

As noted above, English presents velarized or dark laterals in word-final positions. For both speakers of English, this was manifest in a raised F1 (corresponding to jaw height) and a perceptibly lowered F2 (corresponding to posterior or retracted tongue position). Velarization also resulted in substantial dampening of higher frequency energy. For both speakers, all instantiations of [l] or [ɫ] presented clear, dampened, vowel-like formants below 4500 Hz. The first three formants were without exception below 4000 Hz, most often under 3300 Hz.

Subject E1 showed relative regularity in the acoustic quantity or formant values of laterals. In all cases, F2 and F3 were at higher frequencies in word-initial position; correspondingly, F1 was elevated in word-final position. For this speaker, average formant values are calculated as 438 Hz, 1250 Hz, and 2913 Hz for tokens in word-initial position; as 483 Hz, 1183 Hz, and 3117 Hz for those in intervocalic position; and as 806 Hz, 1110 Hz, and 3050 Hz for those in word-final position. Figures 3.25.a through 3.25.c present contrastive spectra of E1 [l] instantiations.

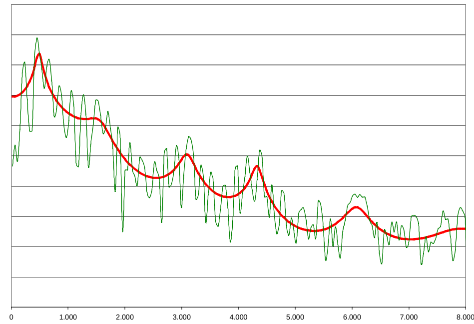


Figure 3.25.a. FFT (green) and LPC (red) spectra of E1 token *late* [laʲt].

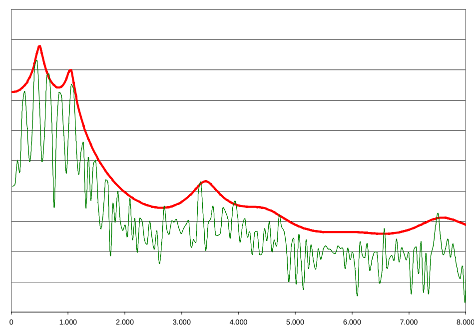


Figure 3.25.b. FFT (green) and LPC (red) spectra of E1 token *mauling* [ma:liŋ].

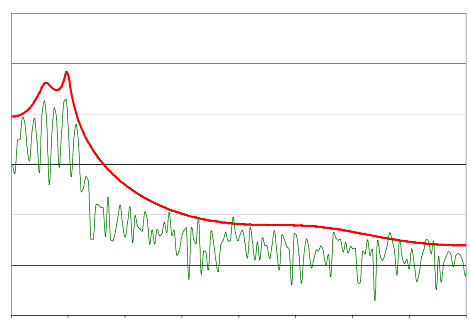


Figure 3.25.c. FFT (green) and LPC (red) spectra of E1 token *awl* [a:l].

Somewhat different formant patterns were observed in speaker E2's instantiations of the lateral. While both English speakers produced velarized laterals in word final positions, E2 tokens did not show higher F2 and F3 in word-initial instantiations. Rather, F3 was higher in word-final position environments. For this speaker, average formants were observed at 538 Hz, 1413 Hz, and 3263 Hz in word-initial tokens; at 483 Hz, 1350 Hz, and 3617 Hz in intervocalic tokens; and at 890 Hz, 1170 Hz, and 3200 Hz in word-final tokens. Figures 3.26.a. through 3.26.c present environmentally contrastive spectra of [l] for speaker E2.

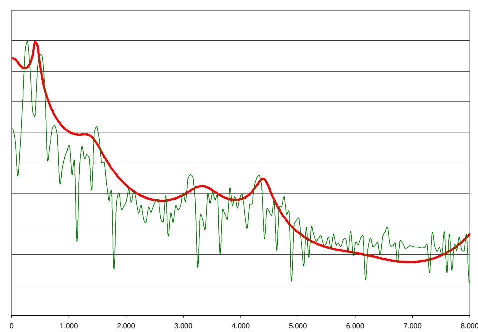


Figure 3.26.a. FFT (green) and LPC (red) spectra of E2 token *late* [la^{j}t].

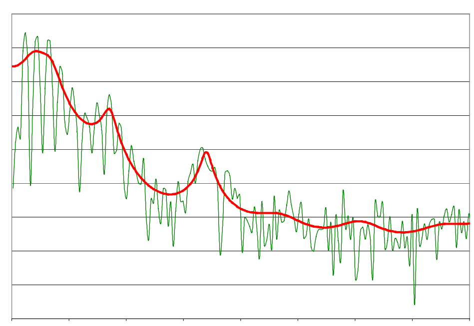


Figure 3.26.b. FFT (green) and LPC (red) spectra of E2 token *mauling* [ma:liŋ].

Figure 3.26.c. FFT (green) and LPC (red) spectra of E2 token *awl* [a:l].

3.3.4.2. Dutch

Like English, Dutch word-final laterals are velarized without exception, this to a greater extent than in the former language. In opposition to English, the only acoustic characteristic of this velarization is a lowered F2 (lower than English), suggesting that velarization is accomplished more by tongue position than by jaw aperture. For both Dutch speakers, word-final articulations also resulted in significantly raised F3.

Speaker N3 displayed little variation in F1 in any of the target environments. For this speaker, the greatest amount of variation was seen in F3, which was consistently lower in word-initial, greater in word-final, and intermediate in intervocalic environments. Average lateral formant values for speaker N3 were calculated at 525 Hz, 1527 Hz, and 2900 Hz for word-initial segments; at 500 Hz, 1433 Hz, and 3083 Hz for intervocalic segments; and at 463 Hz, 1000 Hz, and 3425 Hz for word-final segments. Spectral profiles for this speaker are given in Figures 3.27.a through c.

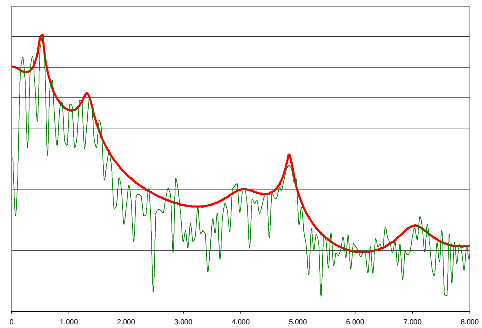


Figure 3.27.a. FFT (green) and LPC (red) spectra of N3 token *laat* [la:t].

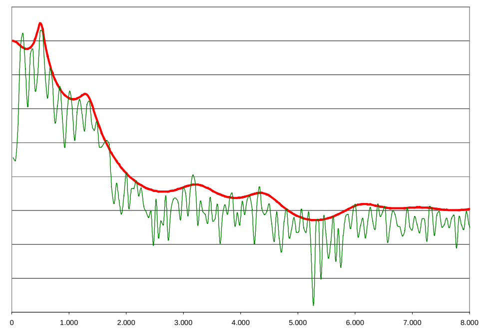


Figure 3.27.b. FFT (green) and LPC (red) spectra of N3 token *halen* [ha:lən].

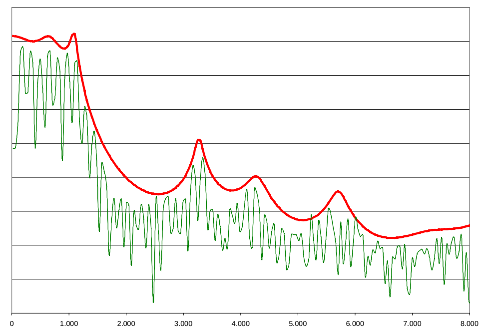


Figure 3.27.c. FFT (green) and LPC (red) spectra of N3 token *haal* [ha:l].

The sole quantitative similarity in speaker N4 was the F2 lowering in word-final environments. For this speaker, F1 was considerably higher in word-initial and intervocalic articulations, although F1 was also measured at ±500 Hz in the latter environments. This formant varied considerably in intervocalic environments, observed between 450 and 1080 Hz. Another quantitative particularity of N4 tokens is F3, which is fairly regular for both word-initial and word-final articulations, but significantly lower between vowels; this may be due to an undershoot effect. Average formant values for laterals articulated by N4 were calculated at 417 Hz, 1683 Hz, and 3833 Hz for word-initial tokens; at 383 Hz, 1467 Hz, and 3000 Hz for intervocalic tokens; and at 413 Hz, 825 Hz, and 3525 Hz for word-final tokens. Figures 3.28.a, b, and c present spectra from N4 lateral tokens.

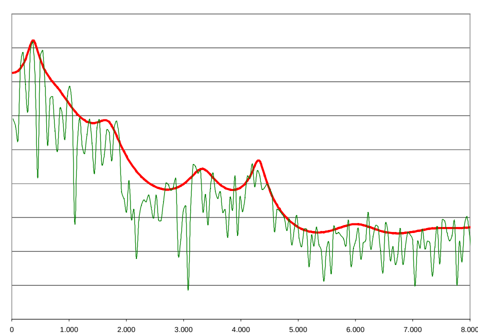


Figure 3.28.a. FFT (green) and LPC (red) spectra of N4 token *laat* [la:t].

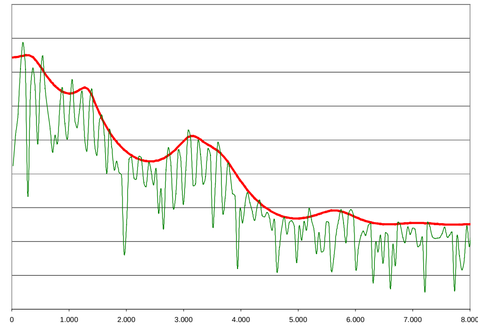


Figure 3.28.b. FFT (green) and LPC (red) spectra of N4 token *halen* [ha:lən].

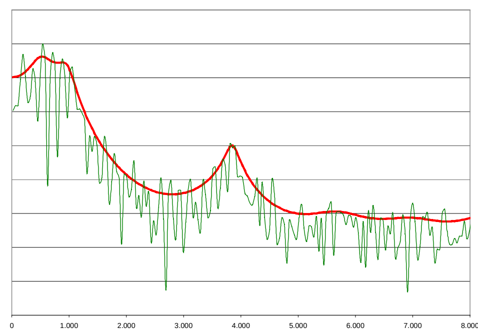


Figure 3.28.c. FFT (green) and LPC (red) spectra of N4 token *haal* [ha:l].

3.3.4.3. French

As mentioned above, a large degree of vowel-to-consonant coarticulatory influence was observed in French tokens. As with other continuants, coarticulation had a greater effect on the output of speaker F5. For both speakers, second and third formants adjacent to [i] showed the most coarticulation and were considerably higher. Contrast is noted in [u] adjacent articulations, when F2 and F3 are lower. No observation of velarization were made, as characterized by either F1 raising or F2 lowering.

Speaker F5 showed substantially more coarticulatory variation than did her compatriot. F1 for this speaker was lowest in word-initial and –final environments; raised F1 in intervocalic environments is likely attributed to VCV transition and/or undershoot. Average formant values for F5 were calculated at 413 Hz, 1813 Hz, and 3350 Hz in word-initial environments; at 388 Hz, 1750 Hz, and 3263 Hz in intervocalic environments; and at 350 Hz, 2238 Hz, and 3575 Hz in word-final environments. Example spectra from F5 tokens are provided in Figures 3.29 a through c.

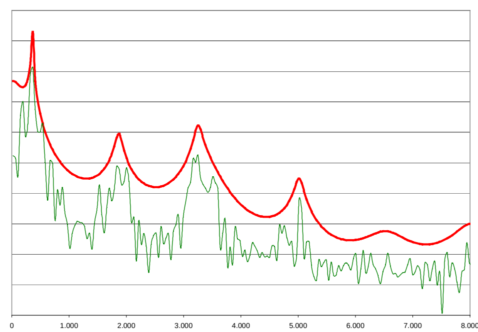


Figure 3.29.a. FFT (green) and LPC (red) spectra of F5 token *la* [la].

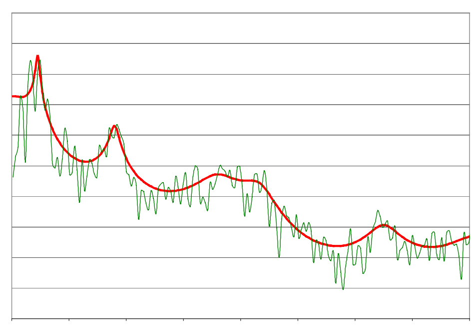


Figure 3.29.b. FFT (green) and LPC (red) spectra of F5 token *alla* [ala].

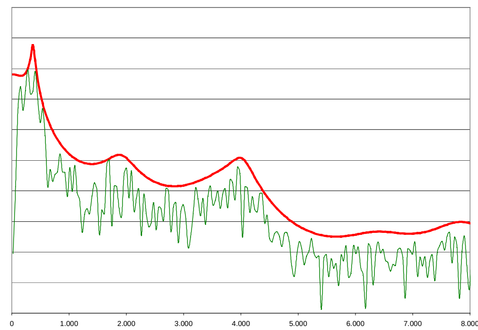


Figure 3.29.c. FFT (green) and LPC (red) spectra of F5 token *balle* [bal].

While speaker F6 displayed the same vowel-to-consonant coarticulation seen in F5 tokens, this was less quantitatively important. Neither F1, nor any other formants showed the effects of undershoot seen in F5. Indeed, formant variation was much more restrained in all aspects for F5. Average formant values for speaker F6 were 450 Hz, 1888 Hz, and 3525 Hz in word-initial position; 350 Hz, 1883 Hz, and 3583 Hz in intervocalic position; and 395 Hz, 2300 Hz, and 3875 Hz in word-final position. Sample spectra of F6 laterals are provided in Figures 3.30.a, b, and c.

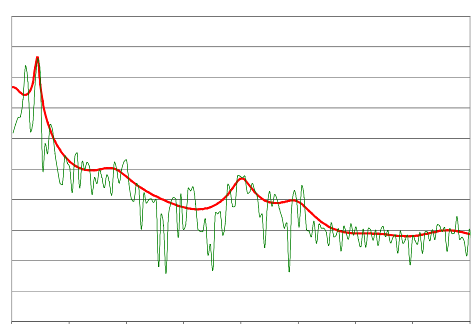


Figure 3.30.a. FFT (green) and LPC (red) spectra of F6 token *la* [la].

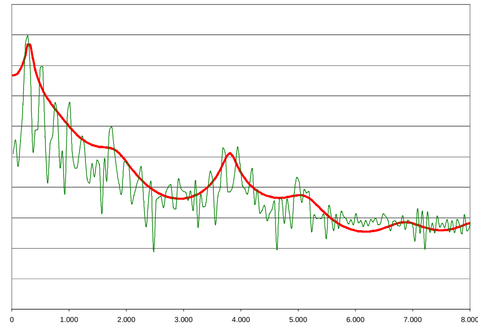


Figure 3.30.b. FFT (green) and LPC (red) spectra of F6 token *alla* [ala].

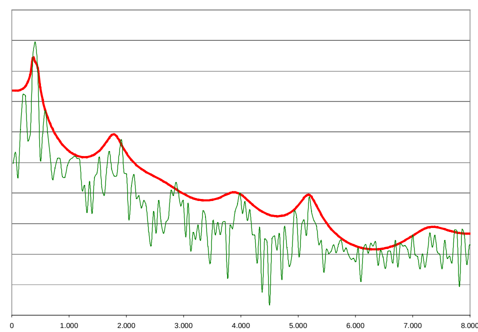


Figure 3.30.c. FFT (green) and LPC (red) spectra of F6 token *balle* [bal].

3.3.4.4. Gestural account

The variable articulatory nature of laterals is acknowledged in phonetic literature, most notably in Ladefoged (1971) and Ladefoged, Cochran and Disner (1977). In the data set constituting this experiment, however, only one lateral with two variables is present, [l]—seen in all of the subject languages—and [ɫ], a positional variant in English and Dutch. All accounts of the "plain" or "clear" [l] attribute the lateral gesture to the tongue tip in contact with the alveolar ridge. The photographic and x-ray evidence of Wängler (1976) supports an alveo-dental articulation in German, for his subject at the very least, where contact is made by the tongue tip and not, as is generally thought, with the blade. Bladon and Nolan (1977) characterize the production of [l] in English using their tripartite index seen for the sibilant fricative [s] in Figure 3.9: [l] is indexed as 1-3, indicating blade and tongue tip heights, respectively. Tranel (1987) makes no distinction between tongue blade and tip in his gestural description of the French [l], merely noting that the tongue tip and front are involved in the articulation of clear laterals.

Dark or velarized laterals are described as involving the bunching or raising of the tongue dorsum—simultaneous to tip-alveolar closure—in Gimson (1970). This production results in a secondary acoustic characteristic, a resonance chamber in the oral cavity in the region of the velum. More recent studies using both magnetic resonance imaging (MRI) and EPG measurements support previous studies. Narayanan, Alwan and Haker (1995) note a lack of full occlusion for [ɫ] in their American English subjects. Here, lateral airflow is produced by tongue bunching, i.e. constriction of the dorsum. Describing dark [ɫ] in both Dutch and English, Collins and Mees (1996) provide a similar description for the latter language, describing the resultant secondary resonance as "[ʊ]-type" (169). In vocal tract drawings, such as those presented in Figure 3.31, they attribute the differences between the two languages to a lack of complete alveolar closure in Dutch. According to their description, the bunching of the tongue dorsum and correspondent raising of the tongue front result in lateral passage of air through the buccal cavity. Nooteboom and Cohen (1995) make no mention of this type of articulation. It is unclear whether all Dutch speakers faithfully reproduced this gesture, most certainly the two subjects who participated in this study.

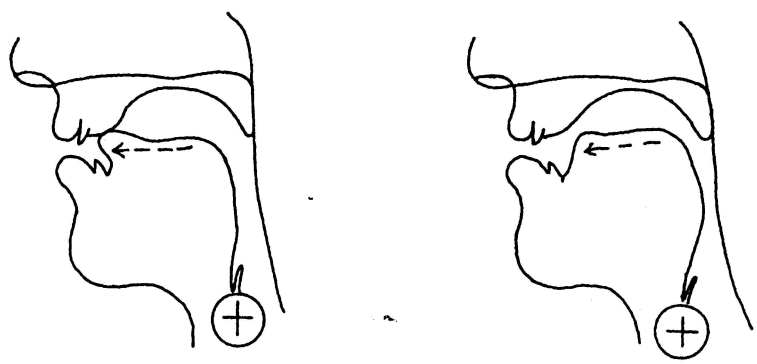


Figure 3.31. Gestural representations of clear and dark [l], respectively, in Dutch (Collins and Mees 1996: 169).

Considerations of secondary articulation and variation aside, a gestural account of the laterals produced by speakers in this study can provide a general account of [l] production. The primary articulator involved in the lateral gesture is the tongue tip. Because our general paradigm does not take laminal versus apical articulations into account, as such contrast always offers a redundant distinction, this is grouped with other gestures involving the tongue front. During movement, the tongue is flat, i.e. non-constricted, save for the muscles catalyzing distal displacement toward the target. The target of articulator movement is the alveolar ridge and adjacent areas. The gesture is non-periodic, as no continual open-close movement of the articulator is implied or observed.

3.3.5. ENGLISH DENTAL (INTERDENTAL) FRICATIVE [ð, θ]

The interdental (or dental) fricatives [ð] and [θ] of English are not present in the consonant inventories of French or Dutch. Voice opposition is maintained in all positions in Standard American English, although there are some instances of free variation and devoicing is a frequent occurrence in word-final positions. Not surprisingly, both dental fricatives are similar in acoustic quality and quantity to [v, f].[16] Like the labial fricatives, dentals present flat spectral profiles consisting of several short formant peaks. For this reason, a range of spectral intensity—and no specific formant value—is given for [ð, θ]. Both English speakers produced very short, almost burst-like fricatives in intervocalic environments. The only important quantitative distinction between English speakers was a slight difference in the lower overall spectra of E1 in relation to those of E2. Figures 3.32.a and b offer samples of interdental spectral profiles.

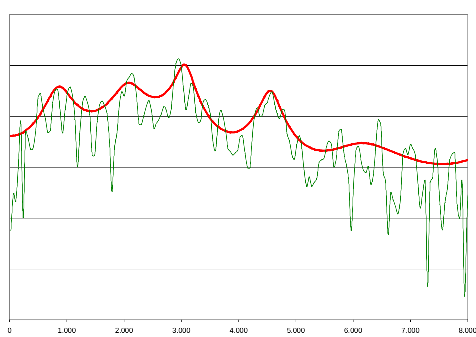


Figure 3.32.a. Example FFT (green) and LPC (red) spectra of [ð]: E2 token *that* [ðæt].

16 A perceptual study of [θ] and [f] by LaRiviere et al. demonstrates that listeners regularly confuse the two segments.

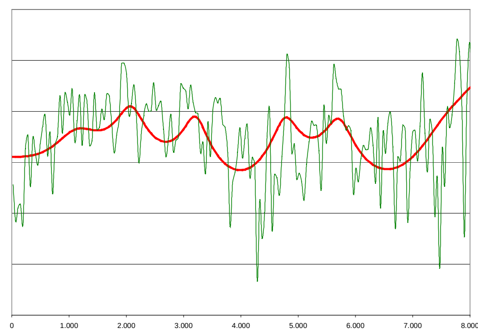


Figure 3.32.b. Example FFT (green) and LPC (red) spectra of [θ]: E1 token *moth* [mɑθ].

Tokens produced by subject E1 resulted in regions of spectral prominence ranging from 2833 Hz to 6183 Hz in word-initial environments; from 2850 Hz to 6383 Hz in intervocalic environments; and from 2800 Hz to 6200 Hz in word-final environments. E2 tokens resulted in spectral profiles where relative intensity was observed at 3233-6250 Hz for word-initial instantiations; 2983-6033 Hz for intervocalic instantiations; and 2933-6217 Hz for word-final instantiations. Figures 3.33 and 3.34 give averaged spectra of [ð] for both speakers.

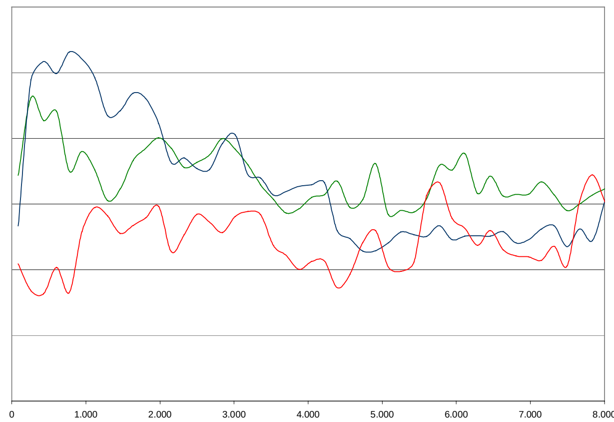


Figure 3.33. Averaged spectrum of labial fricative [ð], speaker E1: *that* [ðæt] (green); *slather* [slæðɹ] (blue); *moth* [ma:θ] (red).

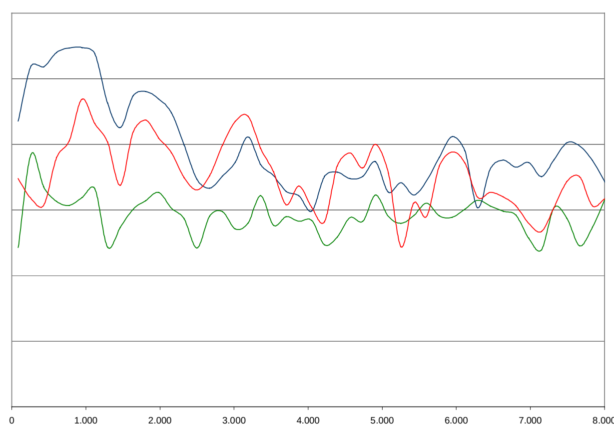


Figure 3.34. Averaged spectrum of labial fricative [ð], speaker E1: *that* [ðæt] (green); *slather* [slæðɹ] (blue); *moth* [mɑ:θ] (red).

3.3.5.1. Gestural account

Surprisingly little gestural description of [ð, θ] is available in linguistic literature. Dental fricatives are a relative rarity in Indo-European: the only other languages within this family containing these segments are Spanish, Icelandic, Danish (intervocalic only), and Greek. Jones (1922) describes dental or interdental fricatives as the result of the tongue tip in contact with the upper front incisors. Gimson (1970) further states that fricative noise is the product of the close aperture between the front teeth or, in the case of some speakers, of the opening created by the tongue protruding between the upper and lower front incisors. Pétursson (1971) and (1974) present some of the most complete description of the inter-dental or dental fricatives. His studies focus on the Icelandic [ð], similar to the voiced segment in English, and supports Jones' and Gimson's accounts. Radiographic evidence suggests a close aperture just behind the upper front incisors, where the tongue front is raised to a position roughly normal to these and the tongue body is uninvolved in consonant articulation (1971: 206-7). Tracings from Pétursson's radiograms are provided here as Figure 3.35.

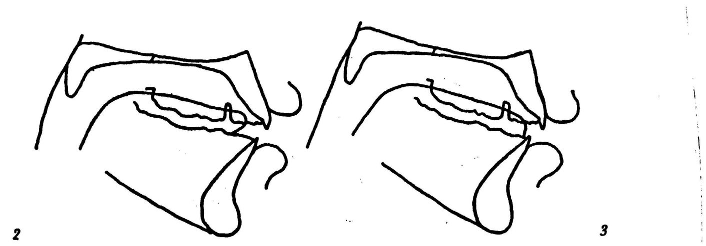


Figure 3.35. Gestural representation of Icelandic [ð] (Pétursson 1971: 207)

Given these accounts of dental fricative articulation, the gesture can be described as involving the tongue front as a primary articulator, whose target is the upper front incisors. The tongue shape is flat, as no dorsal bunching or raising is noted in any study, and there is no periodicity implied by the articulator-to-target movement.

3.3.6. DUTCH VELAR FRICATIVE [x]

The velar fricative [x] of Dutch represents one of the most unique segments included in this study, whether considered in terms of the quality or quantity of acoustic output. This segment, as well as the voiced counterpart [ɣ] of Flemish, are related to [g] of Germanic, and represent a historical case of lenition. Like all Dutch fricatives, [x] is obligatorily devoiced in word-final environments. Despite a history of voiced [ɣ], the segment is voiceless in word-initial positions and intervocalic tokens show a marked devoicing tendency. It is not clear whether intervocalic voicing is intrinsic to the segment, or is merely a phonotactic byproduct.

The most salient qualitative feature of [x] is the widely dispersed, intense quality of sound at frequencies close to those of vowels. Distributed, fricative noise is noted from ±500 Hz to over 6000 Hz, sometimes extending to over 8000 Hz. The [x] of both Dutch speakers displayed anticipatory vowel coarticulation, i.e. interference of the following vowel. This may be attributed to the velar fricative's similarities to [h], a segment almost always coarticulated with a following vowel. Precise degrees of and differences in this interference are discussed with regard to each speaker. Figure 3.36 gives sample spectra of [x].

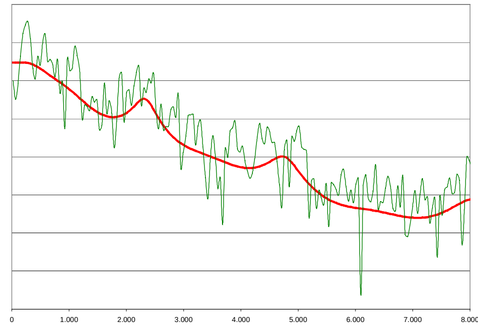


Figure 3.36. Example FFT (green) and LPC (red) spectra of [x]: N3 token *gaaf* [xa:f].

Speaker N3, from Amsterdam, produced [x] having intense, fricative energy from the voicebar extending to ±6000 Hz. The greatest degree of coarticulation is noted in the first formant; other formants vary considerably less with regard to adjacent vowels. Tautosyllabic vowels having a low F2 ([u], [o], and [a]) resulted in a lower fricative F1. Conversely, those vowels with a high F2 (in this experiment, [i]) resulted in a raised first consonant formant. The second formant, ranging from ±3000-4000 Hz, varies considerably more in intervocalic environments. Across vocalic environments, [x] of N3 resulted in average formant values of 1767 Hz, 4083 Hz, and 5250 Hz in word-initial tokens; 2283 Hz, 3550 Hz, and 5100 Hz in intervocalic tokens; and 1683 Hz, 4208 Hz, and 5425 Hz in word-final tokens. Figure 3.37 presents the averaged spectra of [x], as articulated in the three target environments.

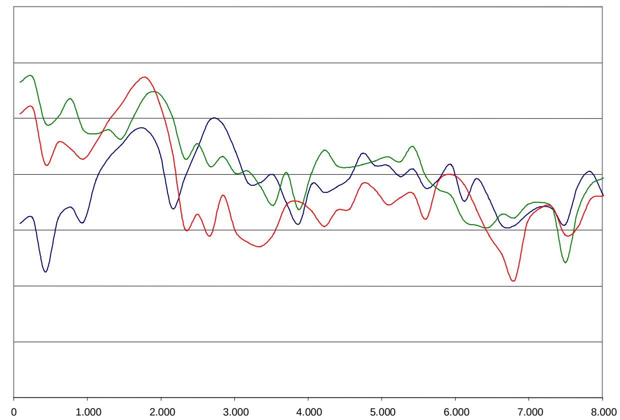


Figure 3.37. Averaged spectrum of labial fricative [x], speaker N3: *gaaf* [xa:f] (green); *agaat* [axɑ:t] (blue); *vaag* [va:x] (red).

N4 is a native of Brabant, where a *zachte-g* or "soft-g" comprises a distinct dialectal variant. This sound is considered by many Dutch speakers to be softer than the "hard" fricative of Holland and the coastal provinces, such as that of N3. This impression corresponds to the quality of sound seen in N4 tokens. These show little of the robust, intense fricative energy of N3 tokens, even though they are equally dispersed and distributed over similar frequencies. Other than intensity, the main distinction of N4's [x] is the degree of vowel-to-consonant coarticulation. In all environments—even word-final—adjacent vowels had a considerable effect on all fricative formants. All formants are lower when adjacent to lower-frequency vowels ([o] and, especially, [u]) and higher when adjacent to higher-frequency vowels ([e] and [i]). Across stimulus environments, average formant calculations for the soft [x] were 2517 Hz, 4667 Hz, and 6583 Hz for word-initial tokens; 2500 Hz, 4367 Hz, and 6250 Hz for intervocalic tokens; and 2092 Hz, 4083 Hz, and 5450 Hz for word-final tokens. Figure 3.38 presents a contrastive averaged spectrum of [x] for speaker N4.

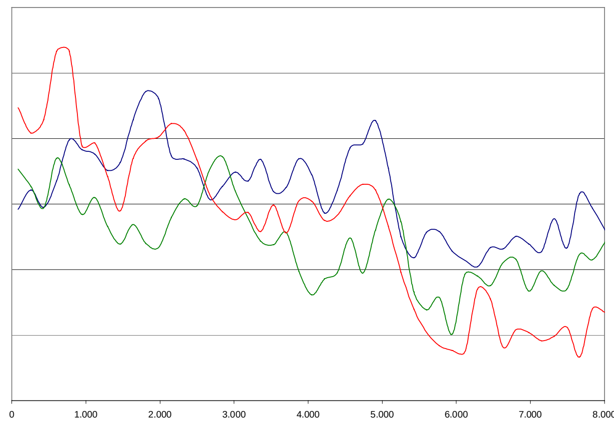


Figure 3.38. Averaged spectrum of labial fricative [x], speaker N4: *gaaf* [xa:f] (green); *agaat* [axɑ:t] (blue); *vaag* [va:x] (red).

3.3.6.1. Gestural account

The fricative noise of [x] is the result of air passage through a reduced aperture at the back of the oral cavity, in the region of the velum. Tracings of Nooteboom and Cohen (1995), based in part on the work of Kaiser (1958) show that the tongue body is raised in the post-palatal region. X-ray evidence in Wängler (1976), in this case for the dialectal German [x], supports the Dutch evidence. Here, the tongue dorsum is tensed and raised, coming in close contact to the velum and, to a lesser extent, the uvular and laryngeal wall. The jaw of Wängler's subject is considerably lower than in the Dutch studies, although no precise measurements are given. Collins and Mees (1996) offer further tracings of the vocal tract configuration for Dutch [x] and for the Southern Dutch [x, ɣ], which they state is closer in articulation to [ç]. Their tracings show more closure than do either Nooteboom or Cohen or Wängler. An illustration is provided in Figure 3.39.

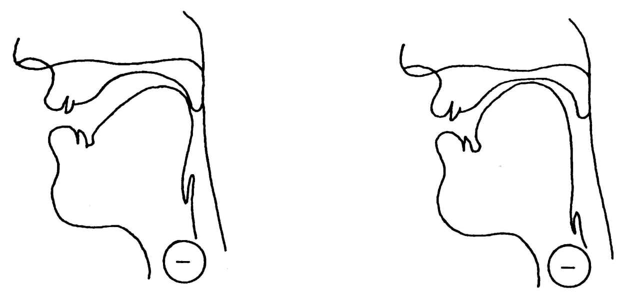


Figure 3.39. Gestural representation of [x] and [ç], respectively, in Northern and Southern Dutch (Collins and Mees 1996: 192).

The formal gestural specification of [x] provides for the tongue body or dorsum as the primary articulator. The target of articulator displacement is the velum, with the tongue shape being flat or non-contracted, save for the muscles necessary for articulatory catalysis. There is no periodicity involved in [x], in contrast to [ʀ], which implies the trilled or flapping movement of the uvula in contact with the tongue body.

3.3.7. English Approximant Rhotic [ɹ] and Rhoticized Vowels

Both American English speakers produced tokens consisting of [ɹ] and rhoticized vowels. The latter were produced irregularly and are heavily influenced by the following rhotic; in these cases, the rhotic itself constitutes the syllable peak, normally reserved for the vowel. This pattern of rhotic use is characteristic of Standard American English and contrasts with other forms, specifically those native to Great Britain, Australia, and New Zealand, as well as the coastal regions of the Southeastern United States and New England. Both experimental subjects produced the target rhotic in all environments, although important quantitative distinctions separate word-initial instantiations from those in intervocalic and word-final environments.

The quality of rhotic tokens produced by both English speakers consisted of vowel-like, dampened formants that show little of the dispersion observed of English fricatives. All significant energy was below 3000 Hz, with any above that frequency being very dampened. Quantitative analysis of [ɹ] shows formants at approximately 500 Hz, 1200-1500 Hz, and 2000 Hz. Two instances of variation were noted in the data, a slightly higher F1 in word-final tokens and a significantly lower F2 in word-initial tokens.

Quantitative analysis of formants for speaker E1 reveals average formant values of 417 Hz, 1133 Hz, and 1850 Hz in word-initial environments; 500 Hz, 1550 Hz, and 1767 Hz in intervocalic environments; and 558 Hz, 1450 Hz, and 2042 Hz in word-final environments. Example spectra for this speaker are given in Figures 3.35.a through c. For speaker E2, contrastive average formant values were 383 Hz, 1183 Hz, and 1938 Hz for word-initial tokens; 533 Hz, 1417 Hz, and 1967 Hz for intervocalic tokens; and 540 Hz, 1420 Hz, and 2040 Hz for word-final tokens. Figures 3.40.a, b, and c are speaker E2 spectral examples.

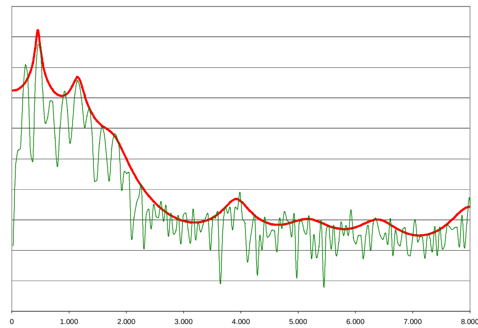


Figure 3.40.a. FFT (green) and LPC (red) spectra of E1 token *rare* [ɹeʲɹ].

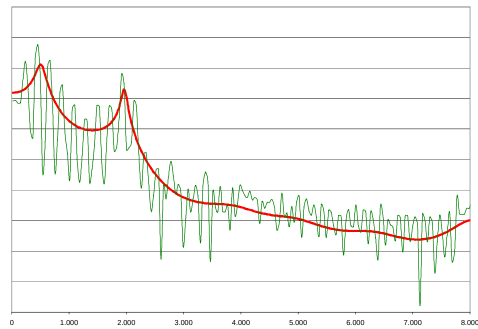


Figure 3.40.b. FFT (green) and LPC (red) spectra of E1 token *daring* [deʲɹɪŋ].

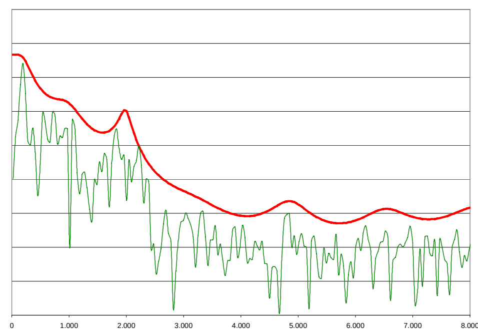


Figure 3.40.c. FFT (green) and LPC (red) spectra of E1 token *car* [kaɹ].

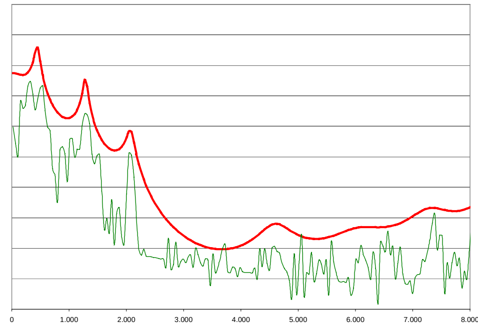


Figure 3.40.a. FFT (green) and LPC (red) spectra of E2 token *rare* [ɹe^{j}ɹ].

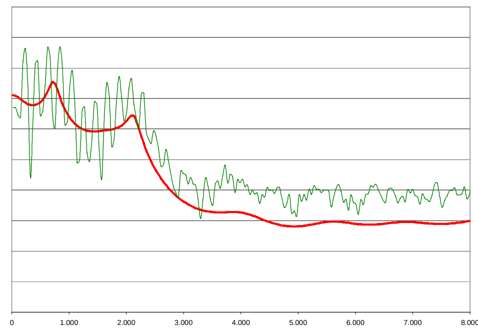


Figure 3.40.b. FFT (green) and LPC (red) spectra of E2 token *daring* [de^{j}ɹɪŋ].

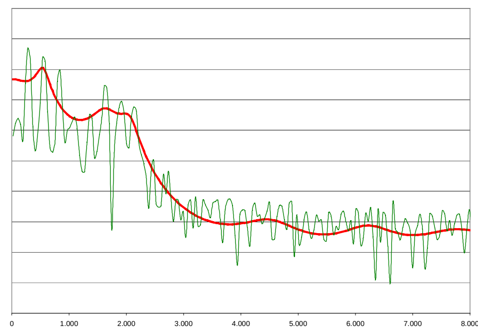


Figure 3.40.c. FFT (green) and LPC (red) spectra of E2 token *car* [kaɹ].

3.3.8. Dutch Rhotic Fricatives and Approximants [ʀ], [ʁ], and [ɹ]

Dutch rhotic tokens display much more variation—both positional and speaker-specific—than do those of English. Significant differences are noted in the rhotic production of speakers N3 and N4, some of which may be attributed to regional and dialectal differences. Discussion of rhotic variation in Dutch is reserved for Chapter Five. Treatment in this section is limited to a description of the phenomenon, including the habits of both speakers, as well as a third speaker whose rhotic is distinct from that of both subjects N3 and N4.

Speaker N3 produced rhotics that were trilled in word-initial and intervocalic positions, and distinctly vowel-like in word-final environments. These correspond to [ʀ] and [ɹ], respectively. The former presented periodic formants as high as 3000 Hz, with additional energy observed as high as 4500 Hz. Trilled instantiations displayed distinct formants nonetheless, qualitatively clearer than those observed of fricatives produced by the speaker. The non-trilled rhotic, seen in word-final positions, is very similar in quality to the American English rhotics presented in 3.3.7. These presented clear, vowel-like formants where all robust spectral energy was concentrated below 3000 Hz. F3 was characteristically lower, at or around 2000 Hz, for the non-trilled rhotic and in relative acoustic proximity to the corresponding F2. Speaker N3 voiced all rhotics in all environments. Figures 3.41.a, b, and c present example spectra from this speaker.

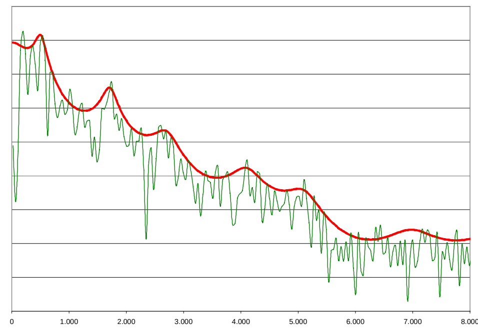


Figure 3.41.a. FFT (green) and LPC (red) spectra of N3 token *raar* [ʀa:ɹ]: the first rhotic is presented.

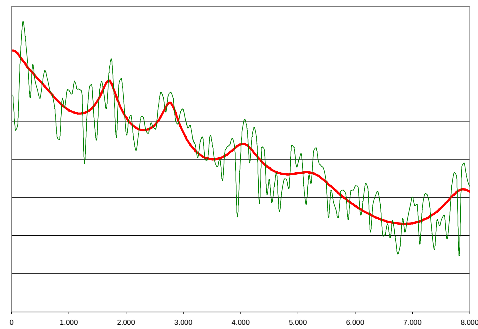


Figure 3.41.b. FFT (green) and LPC (red) spectra of N3 token *hare* [ha:ʀə].

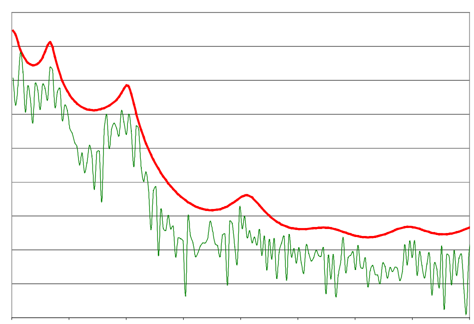


Figure 3.41.c. FFT (green) and LPC (red) spectra of N3 token *haar* [ha:ɹ].

N4 produced rhotics with none of the variation seen in speaker N3. In all environments, tokens from N4 showed dampened, fricative-like formants. As with the trilled variants of N3, all of N4's tokens consisted of energy at or below 4500 Hz, where the most intense acoustic energy was concentrated at approximately 3000 Hz. Tokens of this speaker displayed two other particularities: voicing and intervocalic reduction. In word-final position, instantiations were either voiceless or showed significant devoicing. Intervocalic rhotics where characteristically shorter, even burst like, and consisted typically of one period. Average formant measurements were made for this speaker at 1000 Hz, 1633 Hz, and 3033 Hz in word-initial instantiations; at 900 Hz, 1475 Hz, and 2963 Hz in intervocalic instantiations; and at 863 Hz, 1625 Hz, and 3000 Hz in word-final instantiations. Figures 3.42.a, b, and c provide spectra of N4 rhotics in each of the target environments.

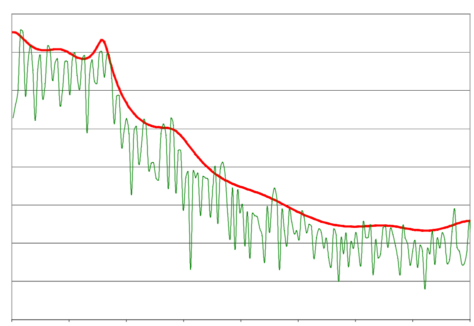


Figure 3.42.a. FFT (green) and LPC (red) spectra of N4 token *raar* [ʁa:ʁ]: the first rhotic is presented.

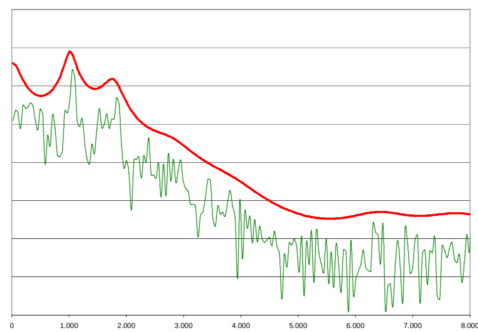


Figure 3.42.b. FFT (green) and LPC (red) spectra of N4 token *hare* [ha:ʁə].

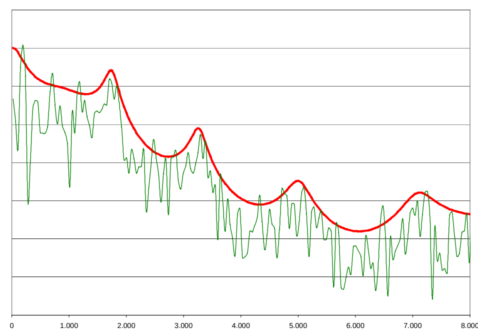


Figure 3.42.c. FFT (green) and LPC (red) spectra of N4 token *haar* [ha:ʁ].

3.3.9. French velar rhotic [ʁ]

The French rhotic presents a number of distinctions and similarities, vis-à-vis the other languages studied here, in terms of both quality and quantity of the output signal. Unlike English and Dutch, little positional variation for either speaker was observed, save for some devoicing in word-final instantiations. As with all other continuants, coarticulation was noted in the rhotics of both French speakers.

In terms of acoustic quality, French tokens are both fricative-like and vowel-like. While formant dispersion is observed—especially in word-initial tokens—this is not as great as with other fricatives in French. Such dispersion is almost completely absent in word-final instantiations and in all cases clear regions of acoustic prominence are seen. The most intense output energy for both speakers was concentrated from approximately 1000 to 3000 Hz, with some energy as high as 5000 Hz.

F5 tokens are notable for their relatively high first formant, as compared to those of speaker F6. Average formant values for F5 are 883 Hz, 1433 Hz, and 3133 Hz for word-initial tokens; 713 Hz, 1563 Hz, and 2950 Hz for intervocalic tokens; and 821 Hz, 1407 Hz, and 3229 Hz for word-final tokens. Sample spectra are given in Figures 3.44.a-c.

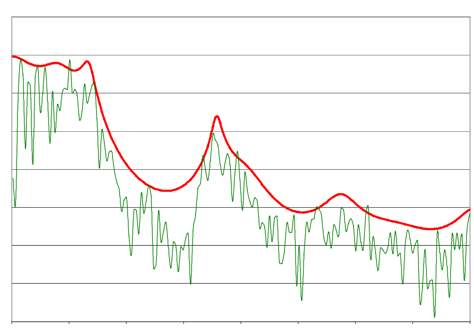


Figure 3.44.a. FFT (green) and LPC (red) spectra of F5 token *rat* [ʁa].

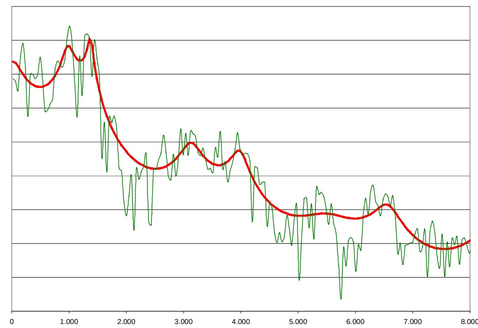


Figure 3.44.b. FFT (green) and LPC (red) spectra of F5 token *arras* [aʁa].

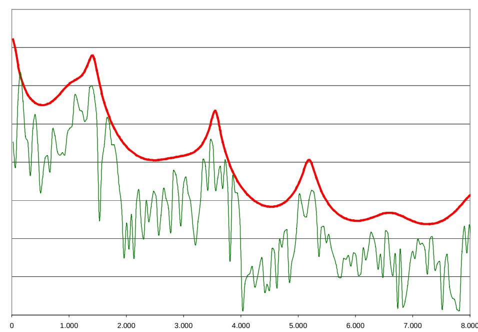


Figure 3.44.c. FFT (green) and LPC (red) spectra of F5 token *par* [paʁ].

Speaker F6 tokens are notable for the positional variance seen in the first formant. These were significantly lower than those of speaker F5 in word-initial and intervocalic environments: word-final instantiations revealed F1 values comparable to the other French speaker. Average formant values were calculated for speaker F6 as 467 Hz, 1450 Hz, and 3167 Hz in word-initial environments; as 650 Hz, 1675 Hz, and 2988 Hz in intervocalic environments; and as 815 Hz, 1557 Hz, and 2972 Hz in word-final environments. Sample spectra for this speaker are given in Figures 3.45.a-c.

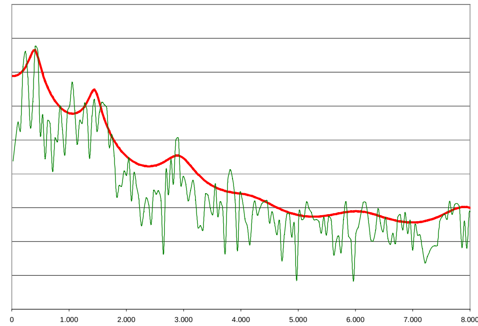


Figure 3.45.a. FFT (green) and LPC (red) spectra of F6 token *rat* [ʁa].

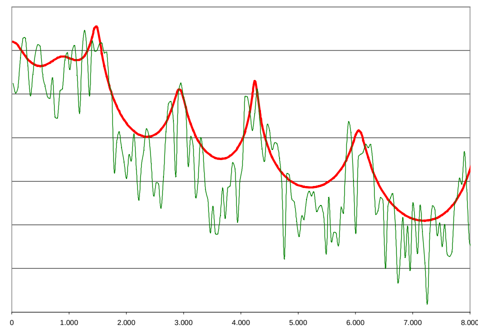


Figure 3.45.b. FFT (green) and LPC (red) spectra of F6 token *arras* [aʁas].

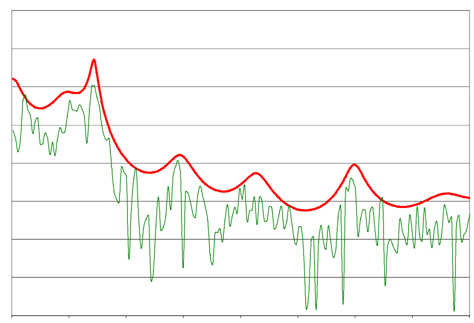


Figure 3.45.c. FFT (green) and LPC (red) spectra of F6 token *par* [paʁ].

3.3.10. GESTURAL ACCOUNTS

There were no fewer than three types of rhotics elicited in this study, [ʁ] of both French subjects and of Dutch speaker N4, [ɹ] of both English subjects and of Dutch speaker N3 in word-final instantiations, and [ʀ] of N3 in word-initial and intervocalic environments.

The approximant [ɹ] is the product of constricted airflow in the center of the buccal cavity, typically in the region of the palate and/or alveolar ridge. In one of the most extensive studies of the American English rhotic, Westbury, Hashi, and Lindstrom (1998) present evidence for the variability of [ɹ] gestures, which they term "r-movements." Their findings show that for the 53 speakers studied, tongue contour during the gesture was nearly always intermediate to the bunched or retroflexed articulations prescribed by previous studies, such as Udall (1958) and Delattre and Freeman (1968). These movements involved the raising of the tongue body, such that a reduced aperture at the palatal or alveopalatal region was created. The researchers noted important coarticulatory effects, mostly evidenced in tongue backing or forwarding, as a result of the phonotactic environment of r-movements. Their results agree with Lindau (1985), who notes that the American English rhotic is accompanied by constrictions in the lower pharynx and at the palate, as well as lip rounding for some speakers. Showing a number of stylized variants of the bunched or retroflexed rhotic, according to tracings by Delattre and Freedman, Figure 3.46 paints a picture—albeit impressionistic—of articulator positions during the rhotic gesture.

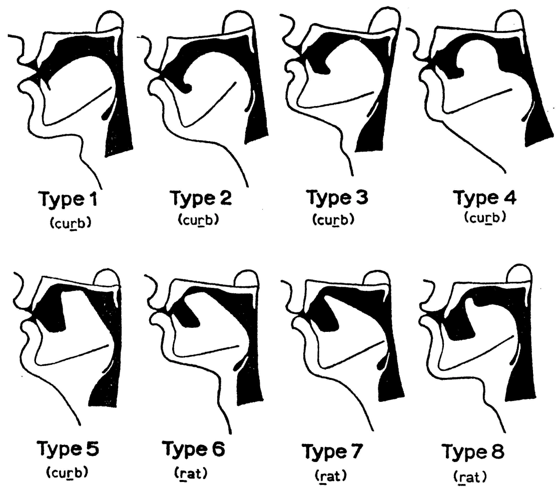


Figure 3.46. Cross-sectional gestural representations of eight types of American English rhotics (Delattre and Freedman 1968).

The [ʁ] of both French subjects and of Dutch subject N4 is produced by close contact of the tongue body with the velum, often in close proximity to the uvula. When egressive lung air passes through the lateral opening at this point of closure, a turbulent airflow is created. Frication is produced as the tongue back is in relatively closer contact to the velar region. By contrast, if the opening is relatively greater, turbulence is lessened and the quality of formants is less distributed and more vowel-like. This is especially noted in word-final instantiations for all speakers having a velar or uvulo-velar rhotic.

Some of the best indirect accounts of [ʁ] production can be gleaned from pedagogical descriptions of this segment, especially for the French [ʁ]. Jourdain and Schuler (1998) advise two methods in their discussion of R-learning in the foreign language classroom. The first of these involves having students produce a “harsh” or fricated [h]. This sound is described as similar to the French rhotic. Once the pharyngeal fricative is mastered, students use the same gesture for [ʁ] in environments where the normative pronunciation was problematic. A second tactic discussed by the authors is referred to as the “gargle technique.” This requires students to mimic the gesture involved in gargling: raising the tongue back and root along the pharyngeal wall, resulting in a constriction at the velum and uvula. Tranel (1987) describes the production of the Standard French rhotic as “a pharyngeal constriction produced in part by drawing back the root of the tongue,” noting that “this constriction may occur in two different places, more toward the lower end of the pharynx or more toward the upper end” (142). The constriction he describes takes place in the velar and uvular areas, such that no distinction between either place of articulation can be advanced. Much of Tranel’s description derives from Delattre (1965), who studied the production characteristics of the French /ʁ/, providing tracings based on x-ray data. These are reproduced in Figure 3.47, although they should be taken as approximant representations, rather than a precise account of the rhotic gesture.

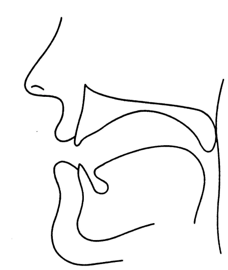


Figure 3.47. Cross sectional gestural representation of the French [ʁ] (Tranel 1987: 144).

The production characteristics of [ʁ] are similar to those implied by [ʀ]. In both cases, the tongue body is raised, causing a constriction in the back of the oral cavity. In the case of [ʀ], this constriction is normal to the uvula, such that the differences in air pressure anterior and posterior to the constriction cause the open-close movement of the uvula, i.e. a trill. Wängler (1976) provides evidence for a simultaneously retracted and raised tongue body in his x-ray study of a German speaker. In this case, the tongue position results in a substantial cavity of high-pressure air in the pharynx, and a larger cavity of lower air pressure in the mouth, anterior to the velum. Delattre (1971) shows that the uvular trill of German and non-standard French speakers is the result of the tongue body being moved back and slightly up, such that the uvula may move periodically. His tracings, as synthesized by Ladefoged and Maddieson, are presented in Figure 3.48.

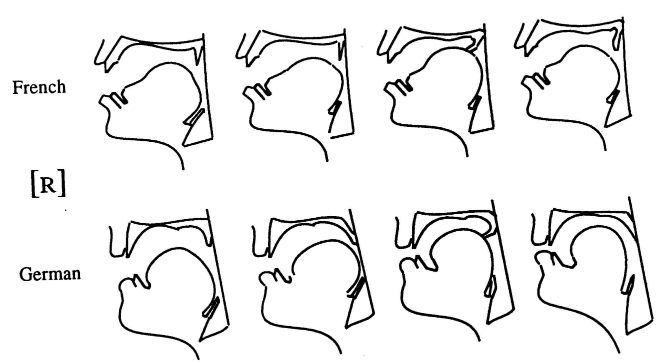


Figure 3.48. Cross-sectional gestural representations of French and German uvular trills (Ladefoged and Maddieson 1996: 229).

3.4. Phonetically grounded phonology

This chapter describes three different continuant consonant inventories and four different r-like sounds in the context of these inventories. Description of continuant inventories and of each constituent of the inventory develops a broad understanding of those gestural and acoustic properties that are significant with regard to segmental distinction within these systems. Tentative examination of the phonetic properties—both articulatory and perceptual—of different rhotics supports the conclusions of Lindau (1985), who asserts that no singular property allows for the union of these segments into a phonetic class or whole.

The goals of this dissertation do not afford phonetics any more than a descriptive role in its development of relational phonology. That relations are grounded in phonetic reality is a given: the functionalist approach denies any assertion of phonological form that is not properly grounded in phonetic reality. However, any systematization of phonetically varying segments and of segmental relations within a larger construct is inevitably and necessarily a phonological enterprise. The quest for emergent categorization, or the "definition-by-opposition" of sound segments, implies the comparison of regularities and the delimitation of significant differences with regard to parametrically related sound segments. Phonetics may at some point be capable of dealing with the variant nature of gestural acts and acoustic realities; until then, phonology provides the only means to explain how such divergent segments may be attributed to a common class or group.

CHAPTER FOUR. RELATIONAL PHONOLOGY: SYSTEM DYNAMICS AND RHOTIC INTEGRITY[17]

4.1. Introduction

The experiment outlined in the previous chapter provided acoustic and articulatory descriptions of the continuant consonant inventories of American English, Dutch, and European French. Using this speaker data, I map these inventories and their respective components allowing for discussion of systemic phonological relations. Analogous use of the term "map" makes reference to the division of entities that are subordinate to a larger whole. Geographic maps divide a landmass, such as a continent, into countries, provinces, and cities. Similar to geographic divisions, any productive whole of a language—in this case, a phonological inventory—may be broken down into component parts. Like states and geopolitical entities, phonological segments define themselves according to their differences as much as they do according to their similarities. Whereas the components of a landmass are sociopolitical entities, determined in part by geography and history, those of the phonological system are articulatory and perceptual entities, determined by gestures and acoustics. Also akin to geographic units, phonological entities—or segments—demonstrate areal and temporal variability. In the case of sociopolitical states or divisions, variability is result of movement of peoples or changes in the sociopolitical landscape.

[17] My use of the term "dynamic" owes much to Banczerowski (1987), who articulates the notion of dynamics in terms of the energy and force of phonetic features, applied to a phonological grammar. While I do not follow his formulae or employ his terminology strictly, this work provides a critical understanding of the logical formulation of dynamic interaction.

Phonological variation is subject to phonetic implementation, as well as sociolinguistic and historical influences.

The analogy drawn between the geography of land and that of the speech system cannot be taken as absolute. Terrestrial reality is grounded in a tactile, non-malleable reality. Any changes in the earth's composition, hence in the possible composition of geographic entities, are restrained by the nature of those entities. The geography of human speech is not entirely plastic and results from the interaction of several dimensions, of which two are the competing drives of articulation and perception. The principles of functionalism outlined in Chapter Two provide for distinct maps of this phonological terrain. The first of these considers the gestural patterns implied in the production of each segment, asking how one segment is distinct from all others with which it is in relation. The second looks to acoustic output and asks what qualitative and quantitative properties are important in the delimitation of one segment from another. The most succinct differential definition, i.e. one that does not overspecify the distinctive characteristics of the segments in question, promotes similarity. My perspective provides for the definition of a segment in dynamic relation (both positive and negative) to its neighbors, i.e. those segments that are co-constituents of the phonological whole in question. The product of this effort may be compared to the theoretically derived system provided in 2.5.

4.2. The gestural-articulatory inventory

In Chapter Three, discussion of the gestures involved in the production of each continuant consonant is limited to include only the broadest description of articulators and articulator movements. Such generalization reflects the functionalist understanding of categorization provided by Boersma (1998) and outlined in Chapter Two, namely that categorization is an emergent property, specific to each system, and that categories (or features) are not a priori available to or present in all systems. For the phonological purposes of this chapter, it is not necessary to look at micro-displacements or at the precise target of articulator movement, for example, as such variations are not understood to be a distinctive property of the gesture itself. To provide just one example, both Westbury et al. 1998 and West 1999 note that production of the American English /r/ is highly variable; each production variant has a basic pattern in common with all others and, furthermore, these variations have little effect on perception (such considerations are taken up in 4.3). My discussion of relational phonology looks only to those properties of the segment that allow it to be distinguished from all other, related segments. Only once articulatory segments are described in terms of their most basic components is it possible to map or provide the relational description of a system.[18]

My reference to gestural components is hardly innovative. A lengthy phonological tradition analyzes phonological units according to these properties.

[18] Flemming (2000) provides a unified model of phonetics and phonology. While not specifically addressing questions of macro-versus-micro gestures, his model accommodates specific phonetic detail without compromising the evaluation of phonological processes.

Browman and Goldstein (1989) provide for the analysis of speech according to tract variables and articulator involvement in an effort to explicate variations in production and integrate these possible variations into a phonological grammar. Browman and Goldstein (1990) and Steriade (1990) provide further use of dynamically defined gestures in an output-biased phonology. In both of these studies, a particular segment is viewed in terms of a multiplex gestural relationship. A given speech task is understood to be the combination of multiple tract and articulator possibilities. The combination of segments leads to the selection of outputs in a derivational manner: certain tract configurations, possible in the input task, are impossible in relation to others and do not appear in the output. Use of the term gesture in this dissertation owes much to the above studies, but differs from them significantly. I provide for a simplified gestural specification and do not concern myself with coordination, nor do I pay particular attention to lower-level gestural variation. Only those gestural components susceptible to distinguish segments one from another are involved in this analysis. While I do not deny the validity of such analyses or models, I propose phonological definitions that are purposefully large enough to encompass all variations seen in the data.

I employ four properties of the gesture, understood as a physical event, to characterize the segments in question in this study. These are the articulator, articulatory target, articulator shape during the phase of the gesture in which the articulator-to-target gesture is made, and periodicity of the gesture implied by the articulator's displacement to or at the articulatory target. The first category describes the catalyst of the articulatory gesture, i.e. the organ positively displaced via contraction of specific muscles for the purposes of speech production. In this study, the relevant articulators are the tongue and lips. Because the tongue is used for all but one of the gestures in this study (/f, v/) and because different groups of tongue muscles are involved in different gestures, distinction between the tongue front and body is motivated. Those gestures involving the tongue front—the blade and tip—are relationally categorized on other grounds, namely the target of articulator movement and the shape of the articulator during this movement. The same is true for gestures implying the tongue body—the dorsum and root. There is no descriptive need of further tongue articulator specification, given the segments that constitute this study.

The second criterion relative to mapping is that of the articulatory target. Here, reference is made to the distal target of the articulator's positive displacement, i.e. the immobile secondary articulator that either comes into contact with the primary articulator or proximal to which a constricted airflow is created between that target and the articulator. For description of the continuant inventories, the range of possible targets is reduced to five categories: the incisors, alveoli, palate, velum, and uvula. The last two targets are confused in the case of French (cf. discussion of [ʁ] in Chapter Three).

The third criterion is that of articulator shape and is relevant only for those gestures involving the tongue. Here, distinction is made between a flat tongue shape, one that is bunched, and one that is grooved. The term "flat" is used for descriptive economy and should not be understood literally. For the purposes of this dissertation, a flat contour corresponds to a tongue position that is neutral, with the exception of the muscles activated for gestural displacement. The resultant contour is not flat, strictly speaking, but does oppose itself to the bunched and grooved tongue contours. The latter involve the restructuring of the tongue for the purpose of creating a reduced aperture through which air passes. Muscle contraction in these cases serves the purpose of articulator-target displacement and of aerodynamic restructuring the buccal cavity.

The final criterion applied to the continuant consonants studied in this experiment is that of periodicity. A period is defined as one or more open-close-open cycles of articulator displacement. Given the systemic components involved in this study, it is only necessary to provide a positive specification for this criterion, as all non-specified elements are assumed to be non-periodic.

	Articulator	Target	Art. Shape	Periodicity
s, z	tongue front	Alveoli	grooved	no
ʃ, ʒ	tongue front	Palate	flat	no
f, v	lower lip	Incisors	---	no
l	tongue front	Alveoli	flat	no
θ, ð	tongue front	Incisors	flat	no
x	tongue body	Velum	flat	no
ʁ	tongue body	velum/uvula	flat	no
R	tongue body	palate, velum, uvula	flat, bunched	yes
ɹ	tongue body	Palate	bunched	no

Table 4.1. Gestural specification for the continuant consonants

Table 4.1 represents the formal specification of each gesture in terms of the four articulatory criteria outlined above. The goal of gestural descriptions such as those provided in the table is to aid in the relational description or mapping of each segment. Figures 4.1 through 4.5 provide schematic representations of five different relationally defined systems examined in Chapter Three. In each of these figures, which are meant only as visual representations and not as physical descriptions of gestures, a segment occupies a specified gestural space. The combined gestural is a delimitation of the entire gestural space afforded the linguistic subsystem of continuant consonants.

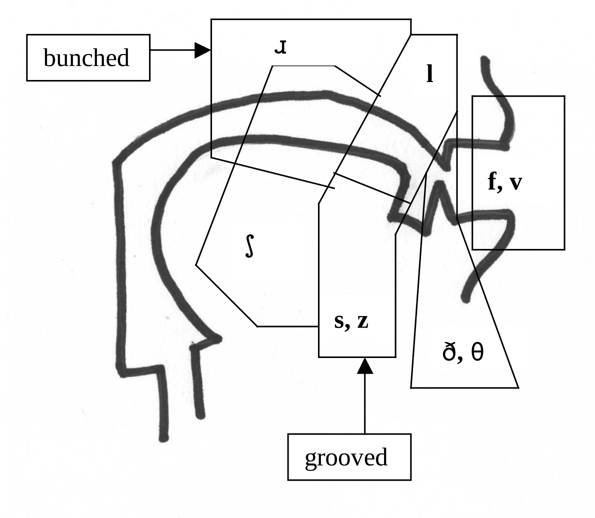


Figure 4.1. Relational description of the gestural system of American English continuant consonant inventory

Conceptualization of the French continuant consonant system {v, z, l, ʒ, ʁ} is distinct from that of English, as presented in Figure 4.2. The same three variables of English are present, although the internal organization of each variable provides for system contrast. The articulator plane is identical to that of

English and is composed of lip, tongue front, and tongue body units. The target plane is relatively more complex, having four components, the lips, alveoli, palate, and the velum/uvula. Note that the latter is considered a singular distal target. By contrast, the articulator target plane is relatively simpler than that of English, consisting of only flat (unmarked) and grooved contours.

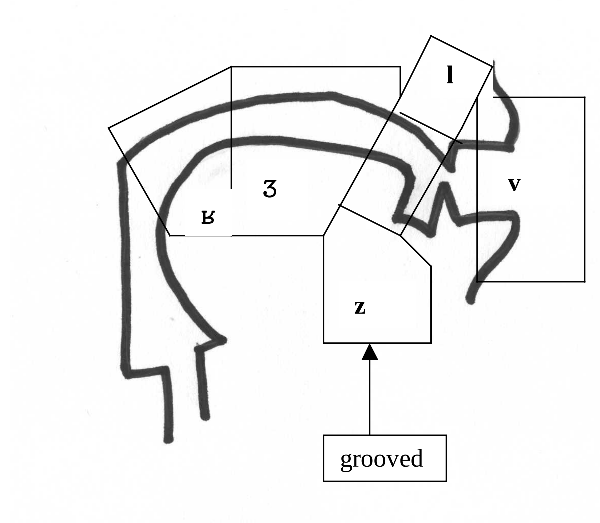


Figure 4.2. Relational description of the gestural system of Standard European French continuant consonant inventory

The type of Dutch spoken by subject N3—which I refer to as Amsterdam Dutch—must be distinguished from that of subjects N4 and N7. This is perhaps the most complex articulatory relation of any investigated in this dissertation, as given in Figure 4.3. The subset of continuant consonants {f/v, s/z, l, ʃ, x, R} is described using virtually all of the planes and components available from Table 4.1. The components of the target plane are the incisors, alveoli, palate, velum, and uvula; those of the articulator plane are the lips, tongue front, and tongue body, and each of the three tongue contours is needed to relationally distinguish sounds from each other. Note that no distinction is made between the two rhotics produced by this speaker in this mapping; discussion of allophony is reserved for Chapter Five.

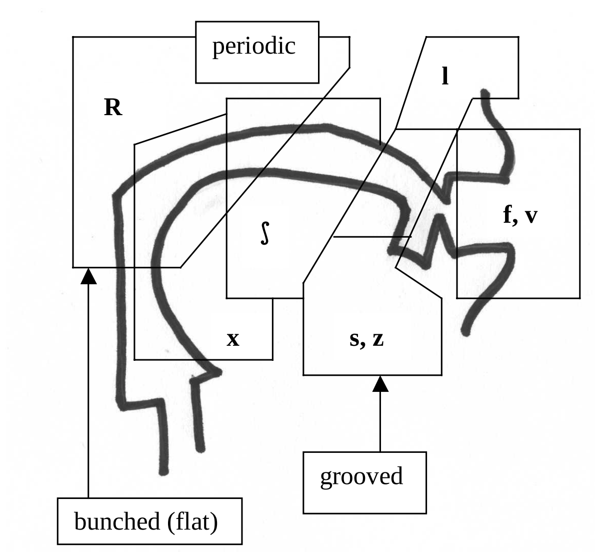


Figure 4.3. Relational description of the gestural system of Amsterdam Dutch continuant consonant inventory

The Dutch system of subject N4—Brabant Dutch—is relatively simpler than that of Amsterdam Dutch. This system is presented in Figure 4.4. In relationally distinguishing the continuant consonant subset {f/v, s/z, l, ʃ, x, ʁ}, the periodicity plane is absent. On the target plane, the fifth component is the velum/uvula, as opposed to the uvula of Amsterdam Dutch. Additionally, there are only two tongue contour components to the articulator shape: flat and grooved.

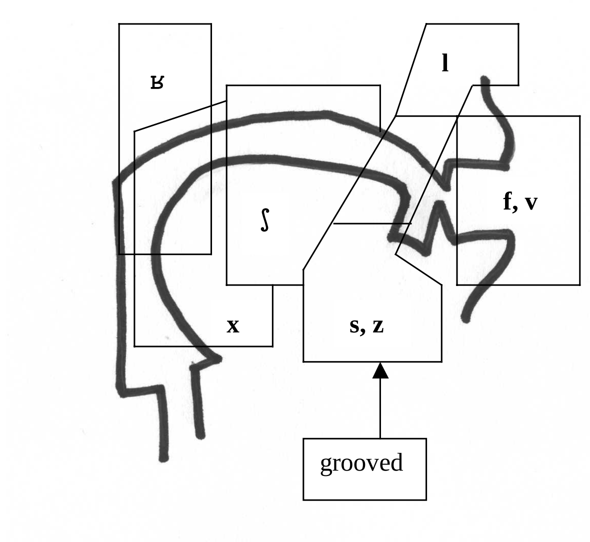


Figure 4.4. Relational description of the gestural system of Brabant Dutch continuant consonant inventory

4.2.1. ARTICULATORY RELATIONS

The introduction to this chapter analogously refers to the complex interaction of the members of a sound system as being a complex gestural geography. 4.2 provides schematic diagrams or maps of the articulatory relations of these members—tentatively titled segments—in each of the systems studied in Chapter Three, showing that each segment, understood as a gestural event, occupies a specific portion of the articulatory space available to the system.

An articulatorily conceived, phonological segment is defined as having specific properties that are themselves components of larger planes, i.e. functionally determined features of the larger systemic context. Given that the planes and their multiplex properties are emergent properties of the systemic context (for simplicity's sake, a grammar), these will logically vary from language to language. It bears reiteration that any attempt to define articulatory relations in terms of a priori categories is a theoretical impossibility within this approach. Articulatory relations are therefore considered to be the dynamic interaction of two or more units, each of which is itself defined by a combination of gestural components, within a systemic context.

Delimitation of the articulatory system is necessary to provide finiteness, i.e. to give boundaries to the larger whole to which the segments are subordinate. In this construction, the system is delimited by its constituent members, and defined in two manners: physiological and productive. Physiological restrictions apply to the size of the system and are the results of human biology; physiological finiteness exists outside the linguistic system. Productive restriction is derived from the need for economy and delimited by the gestural components specified by the units contained in the system. Productive finiteness thus emerges from the segments that make up the whole.

The combinatory result of system dynamics, articulatory integrity refers to the manner by which a given unit of a given system is both alike and distinct from all other units of the same system. For example, the French /ʁ/ is defined by its distinction from /l/ and /ʒ/, by virtue of the target of a flat tongue body displacement, and from /s, z/, by virtue of the same target difference and a different tongue contour. French /ʁ/ is also defined by relation of this segment to /f, v/, distinct by virtue of articulator and target differences. At the same time, /ʁ/ is relationally similar to /l/ and /ʒ/ (flat tongue contour) and to /s, z/ (tongue as articulator). The combination of these similarities and differences defines the articulatory integrity of French /ʁ/.

It is also possible to speak of the patterned articulatory integrity of segments across languages. I refer to this as vertical integrity, that is, the secondary correlation of segments from different systems—themselves defined primarily according to horizontal relation within an organic system—to segments of similar integrity of other systems. Some of these patterns are easily noted. Each language's /l/, for instance, is relationally defined in much the same way across all four systems that comprise the limited scope of this study; specifically, this segment is distinguished from respective /r/ and /s, z/. The rhotics do not pattern so neatly: English /ɹ/ is relationally close to /l/ and /s, z/; Amsterdam Dutch /R/ to /l/, /s, z/, and /x/; Brabant Dutch /ʁ/ to /x/; and French /ʁ/ to nothing. Lindau (1985) noted this complexity in her parametric study, albeit using different terminology (cf. Figure 1.1). The issue of rhotic relational integrity is discussed further in 4.5.1.

4.3. The perceptual inventory

The results of the experiment described in Chapter Three provide acoustic profiles of continuant consonants in each of the target languages. The acoustic output of speakers is defined as the noise energy resulting from articulator movement during pulmonic egression, done for the purpose of communication. This energy is the byproduct of a physical event, much as any noise resulting from biological or mechanical activity. Acoustic energy resulting from a communicative act is, however, different from other noise and any association of perception to acoustic output necessitates systemic categorization, i.e. the attribution of regularity and the limitation of meaningful variation.

Further understood in the term perception is the notion of reparability. This is the possibility that the raw acoustic energy produced by an action or sequence of actions be mapped into an extant systemic pattern or structure and understood by the neuro-auditory system of the perceiver as being congruent to this pattern or structure. Hume and Johnson (2000) provide a preliminary formalization of the interaction of perception and the phonological system. They note that, in addition to the pure mechanics of hearing and of neurological retrieval of stored acoustic patterns, other factors influence perception. These are speech production, linguistic cognition, and social influences on the perceiver. Diehl (2000) notes that human listeners accomplish vowel categorization in a manner similar to relatively simple linear functions, when perceptual units are represented in auditory units such as Barq. This supports the notion that segments must be understood dynamically, rather than as static acoustic or auditory targets, and that phonemes are perceived relationally, rather than absolutely. Previous studies, such as Kingston and Diehl (1995, 1994) relied on features rather than categories. Kingston and Diehl (1994) note that speakers (here, articulators or producers of acoustic signals) anticipate the perceptual effects of their actions (i.e. they use their "phonetic knowledge"). Kingston and Diehl (1995) further distinguishes between the actions and phonetic knowledge of speakers and listeners, in a work specifically described as motivated by a strong auditorist perspective (7-8). Whereas the production of a speaker is linked in part to eventual perception, the activities of a listener are not linked to articulation. In this perspective, covariance can be integrated into intermediate perceptual properties; likewise, one acoustic property may contribute to several perceptual properties. A final study worth mention is Holt and Kluender (2000), who note that such general auditory processes are not specific to speech and that these mechanisms contribute to perceptual accommodation of coarticulation and of variance.

As stated in Chapter Two, discussion of perceptual output in this dissertation does not take into account the auditory processes implied by perceptual activity, nor do I integrate neuro-cognitive mechanisms in the discussion. Data is provided here in the in Hz rather than in Barq and is analyzed from an acoustico-perceptual, rather than auditory-perceptual perspective. Of interest in this section are the acoustic signal and the description of regular differences and similarities within specific systems with regard to these acoustic signals. The scope of this work does not allow for treatment of auditory perceptual mechanisms and for the repair of faulty acoustic signals, however interesting and valuable these questions may be, although it is hoped that future work involving relational phonology will encompass such variables as well as others in its scope.

Use of acoustic or perceptual contrast as a means of explaining phonological behavior is not new to this dissertation, although few have provided perceptual grammars of consonants or consonant inventories. Liljencrants and

Lindblom (1972) are two of the first to apply perceptual contrasts to the definition of a larger inventory—in this case that of vowels—and inspired more recent attempts to define vowel inventories according to the interaction of perceptual properties (cf. Schwartz et al. 1997a and 1997b). Flemming (2001, 1995) also uses acoustic signal contrasts in his discussion of DT and of segmental oppositions; Padgett's analysis of emergent palatalization in Russian (2001) provides an interesting application to liquid consonants. As noted in Chapter Two, primary concern for the discussion of a relational phonology of rhotics is the maintenance of minimum distinction between system members and, in the case of Schwartz et al., the relative focusing of the overall system space. My use of perceptual distinction derives from these works, albeit with a crucial distinction. The complexities of consonant inventories—namely the interaction of perceptual contrasts not seen in vowel inventories—shatters the elegant symmetries favored by Schwartz et al., Flemming, and Liljencrants and Lindblom. Furthermore, the organic definition of consonants that I propose goes much further than Padgett, as it defines the focus segment entirely in reference to other system members.

In order to provide for the relational distinctiveness of perceptual attributes, sound quantity and quality are taken into account. These are introduced in Chapter Three by referring to the type of noise energy and the measurement of its spectral prominence, respectively. Integrating perception into a relational map of consonant inventories requires that each acoustic characteristic be systematized, allowing for the inclusion of variance into a greater structure. For the purposes of this dissertation and given the reduced inventories at hand, only two qualitative distinctions are necessary: the presence of dispersion and of defined formant peaks. These are abbreviated as “dispersion” and “formants” in subsequent charts and discussion.

Each of the qualitative labels is privative (i.e. non-binary) and explicitly linked to the type of energy associated with a sound. Neither label is exclusionary. For instance, [z] presents dispersed energy and relatively distinguishable formant peaks; it is therefore categorized as being dispersed and having formants. The fricative [v] has dispersed energy with relatively undistinguishable formant peaks, and is categorized only as being dispersed. A final example is [l], which is not dispersed, but presents clear, vowel-like formants.

In contrast to qualitative attributes, quantity refers to the location of acoustic energy or regions of spectral prominence. My approach does not capture all noise frequencies in its quantitative analysis, but considers only the most prominent peaks. These correspond to the important acoustic components of a sound segment, i.e. those properties that permit a listener to distinguish it from a range of meaningful sounds. Raw quantitative measurements vary according to diverse factors, such as the position of segments in the word or the coarticulatory interference of adjacent segments. A relational mapping must therefore look at the range of possible quantitative variations observed in speaker data and present these in a manner that includes the greatest number of variations without overstating the phonological system dynamic.

For the purposes of description here, and to better practice a theory of relational integrity, the delimitation of quantitative space is accomplished by taking the average of all measured data points for the formants of a given sound. The range of variation is calculated according to the standard deviation (rounded to the nearest 5 Hz) of all token measurements from the average. Thus, the greatest number of systematic variations is included in the categorized segment quantity, excluding superlative and outlying measurements. Qualitative and quantitative specifications for each of the studied continuants are provided in Table 4.2.

Figures 4.5 through 4.8 are modeled spectral maps of the perceptual systems established by the acoustic output of each speaker. These represent the interaction of segments, noting spatial distinctions and similarities. Each segment is, in its respective system, both in relation to other segments (by virtue of shared quantitative space and qualitative patterns) and distinct from all others (by the particular organization of quantitative and qualitative patterns).

The English perceptual relations presented in Figure 4.5 capture interaction within this perceptual system. Each segment is distinguished by at least one qualitative or quantitative difference. In case of /f, v/ and /ð, θ/, the qualitative resemblance noted in 3.3 is less important for perceptual distinction, given the regular quantitative difference seen in the entire spectral profile. Distinction between /f, v/ and /s, z/ is not based on quantity, but on quality, namely the presence of clear formants. Likewise, /l/ and /ɹ/, the two vowel-like members of the inventory, are distinguished mainly by their respective third formants: F3 of /l/ is greater than that of /ɹ/.

	Quality	Quantity (Hz)
English ɹ	formants	(E1) 390-630, 1090-1700, 1810-2140 (E2) 360-630, 1080-1630, 1805-2175
L	formants	(E1) 420-790, 1055-1295, 2865-3175 (E2) 465-875, 1105-1485, 3090-3560
s, z	dispersion, formants	(E1) 4735-5215, 6075-6405, 7695-8015 (E2) 4590-4860, 6140-6610, 7610-7980
ʃ	dispersion, formants	(E1) 2850-3110, 3875-4205, 5875-6085 (E2) 2975-3155, 3965-4425, 5965-6575
f, v	dispersion	(E1) 4960-8100 (E2) 4475-8045
ð, θ	dispersion	(E1) 2705-6450 (E2) 2890-6380
French ʁ	dispersion, formants	(F5) 645-965, 1220-1690, 2835-3425 (F6) 460-930, 1290-1840, 2800-3240
L	formants	(F5) 340-430, 1480-2390, 2895-3895 (F6) 345-445, 1495-2575, 3220-4120
Z	dispersion, formants	(F5) 4540-5660, 6325-6665, 7710-7970 (F6) 4495-5475, 6070-6450, 7310-7570
ʒ	dispersion, formants	(F5) 2735-3095, 3655-4375, 6070-6410 (F6) 2590-3120, 4030-4200, 6085-6305
f, v	dispersion	(F5) 5015-7905 (F6) 4980-7665
Amsterd. L	formants	425-565, 995-1575, 2870-3430
(N3) R	formants	420-640, 1300-1920, 2065-2935
s, z	dispersion, formants	4185-4995, 5980-6450, 7330-7830
ʃ	dispersion, formants	2775-2925, 3705-4065, 6010-6210
X	dispersion, formants	1405-2305, 3570-4460, 5030-5570
F	dispersion	4115-7890
Brabant ʁ	dispersion, formants	745-1085, 1270-1870, 2790-3200
(N4) l	formants	360-450, 835-1695, 3045-3605
s, z	dispersion, formants	4205-4995, 6000-6430, 7530-7800
ʃ	dispersion, formants	2580-2890, 3605-4175, 5020-6150
X	dispersion, formants	1580-3020, 3420-5180, 4970-6900
F	dispersion	4290-7930

Table 4.2. Perceptual specification for continuant consonants

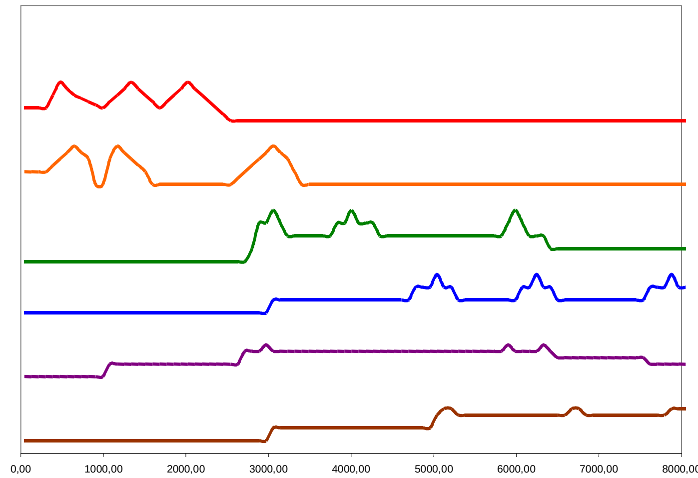


Figure 4.5. Modeled perceptual system of English: /ɹ/ (red), /l/ (orange), /ʃ/ (green), /s, z/ (blue), /ð, θ/ (purple), /f, v/ (brown)

The French perceptual system is not significantly different from that of English with regard to the relational dynamic of the participant segments, as seen in Figure 4.6. The relation of /z/ to /ʒ/ mirrors that of the corresponding segments in English and Dutch; the quantitative position of /v/ is likewise closely related to /z/, where distinction is derived from qualitative considerations. With regard to /l/ and /ʁ/ (the latter segment qualitatively having both fricative dispersion and vowel-like formants), a primary quantitative distinction stems from the lower regions of acoustic energy, as F1 and F2 are quantitatively distinct.

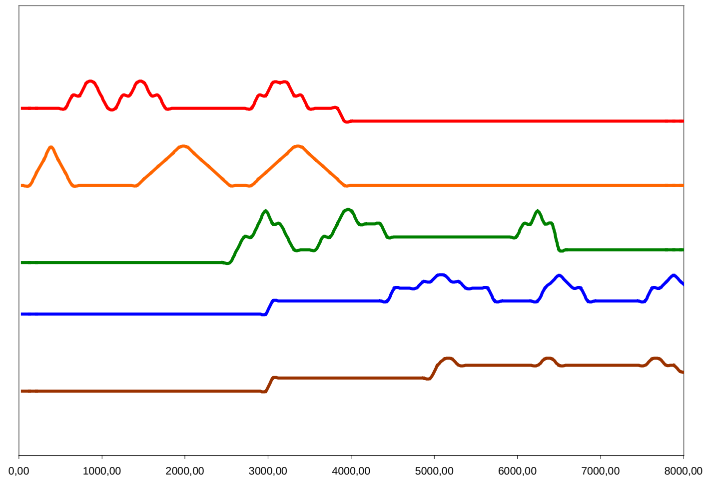


Figure 4.6. Modeled perceptual system of French: /ʁ/ (red), /l/ (orange), /ʒ/ (green), /z/ (blue), /v/ (brown)

Given the different types of rhotics produced by the two speakers of Dutch, it is not surprising that the relational dynamic established by these segments' inclusion in the continuant inventory should be different. Figure 4.7 provides a relational description for Amsterdam Dutch. Note that, for the moment, no distinction is made between the two very distinct rhotics, word-initial and intervocalic [ʀ] and word-final [ɹ]. Without consideration of the different rhotic allophones, quantitative and qualitative overlap of /R/ and /l/ results. Chapter Five examines this congruence and rhotic allophony.

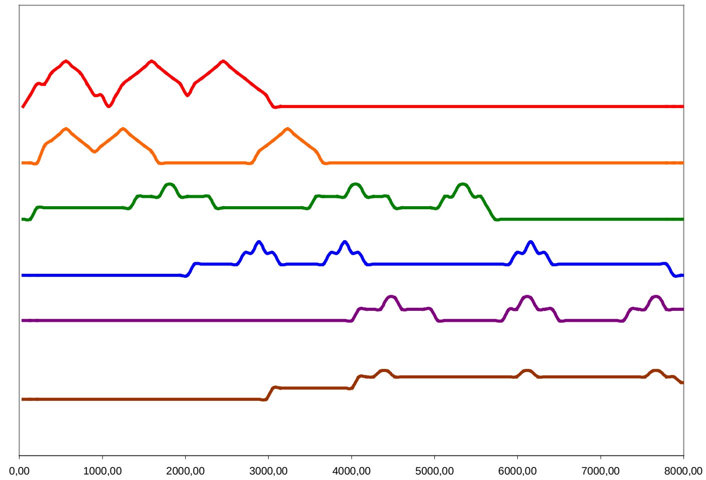


Figure 4.8. Modeled perceptual system of Amsterdam Dutch: /R/ (red), /l/ (orange), /x/ (green), /ʃ/ (blue), /s, z/ (purple), /f, v/ (brown)

A final perceptual mapping is seen in Figure 4.8, the relational description of Brabant Dutch. Here again, there is significant overlap of the rhotic member and /l/, although there are important qualitative differences between the two. With the exception of the lower-frequency formant peaks, a good deal of quantitative overlap of /ʁ/ and /x/ is also present. As is the case for English, French, and Amsterdam Dutch, qualitative factors are most important for the distinction of /s, z/ and /f, v/.

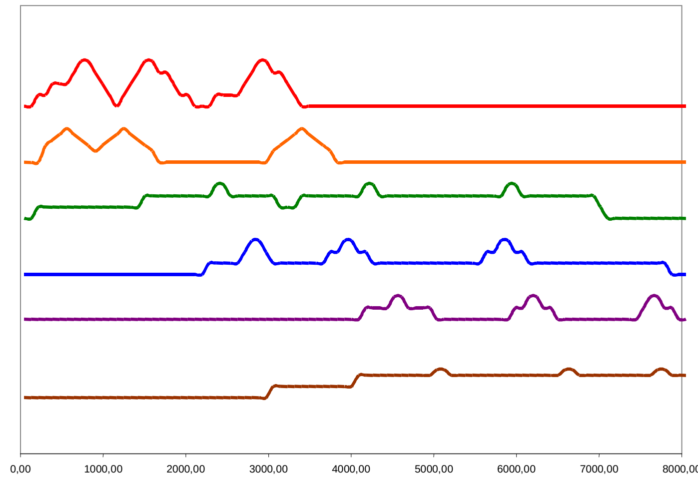


Figure 4.9. Modeled perceptual system of Brabant Dutch: /ʁ/ (red), /l/ (orange), /x/ (green), /ʃ/ (blue), /s, z/ (purple), /f, v/ (brown)

4.3.1. PERCEPTUAL RELATIONS

As in the articulatory representations of 4.2.1, patterns of perceptual relation are noted in the descriptions above. These relations are understood to be the differences and similarities in the perceptual profile of units of a sound system. For example, in all four of the systems presented above, /s, z/ is similar to /ʃ, ʒ/ by virtue of the quality of acoustic energy and to /f, v/ by virtue of its quality. Likewise, /l/ and each of the rhotics present perceptual congruence, both qualitative and quantitative. The ensemble of perceptual differences and similarities within a system provides for a segment's integrity or definition. The perceptual integrity of the French /z/, to present just one example, stems from its qualitative similarities to /ʒ/ and /ʁ/ and its quantitative similarities to /v/, as well as its quantitative and qualitative differences with regard to these respective segments. In essence, /z/ is understood to be similar to and distinct from {ʒ, v, l, ʁ}, and is defined as a distinctly integrated segment by virtue of these relations.

The above definition of the French /z/ is an instance of horizontal perceptual integrity, i.e. the delimitation of French /z/ with regard to the other members of the French continuant inventory. It is also possible to speak of cross-linguistic relational similarities among sibilant fricatives /s, z/. Here, segments from different systems are posited to be similar due to the congruence of relations within the respective systems. For example, the /z/ of French defines itself in much the same way as /s, z/ of English and Dutch. The same case may be made for /l/, /ʃ, ʒ/, and /f, v/, all of which show a good deal of vertical perceptual integrity in the four systems studied here.

The system-specific perceptual integrity of individual rhotics results from different systemic relations. While each rhotic is both similar to and distinct from other members of its respective inventory, these similarities and distinctions are not similar across the four systems. French, English, Amsterdam Dutch, and Brabant Dutch integrate a rhotic segment in a different way, i.e. each is potentially confusable with at least one other member, although this member is not the same in every system. The English rhotic derives its integrity from quantitative distinction from fricatives and from /l/, which shares many /ɹ/ qualities. In French, rhotics are quantitatively similar to /l/, but qualitatively different and relatively distinct from all continuants. Brabant Dutch presents a rhotic only marginally distinct from /l/, depending solely on qualitative considerations for its integrity; this segment is also similar to /x/. Amsterdam

Dutch remains a troubling case: quantitatively and qualitatively, /R/ overlaps both /l/ and /x/, pointing to a more complex relational integrity involving allophony.

4.4. Accounting for systemic relations with violable constraints

In this section, I turn my attention to the formalization of systemic relations, a necessary step in the advancement of any theory of relational phonology and in the construction of relational integrity of segments. Formalization of relations in terms of violable constraints also allows for later discussion of allophony and dialectal variation.

Chapter Two provides articulatory and perceptual grammars of theoretical inventories. Tables 2.1 and 2.2 reflect constraint interaction in the respective grammars, using *HOLD constraints for the articulatory model and MAINTAIN and *MINDIST constraints for the perceptual. In this section, I return to the OT model and describe each language and its respective articulatory and perceptual grammars using a constraint-based model. My definition of the relevant constraints derives from the relation of segments to each other (for the continuant systems) and of the rhotic in relation to the members of its respective system (for these segments, in particular). The reader will note that, in this section as in Chapter Two, use of OT tableaux is purely descriptive and is in no way meant to be considered generative or output-oriented.

4.4.1. ARTICULATORY GRAMMARS

The articulatory drive and resultant articulatory dynamics depend upon the interaction of REUSE and *HOLD. For the specific application of these constraints to articulatory output, their definitions must be somewhat modified from those seen in 2.4.1. The efficiency constraint REUSE is defined locally—for the continuant consonant system—and globally, referring to gestures common to the entire language inventory. These are given as follows:

> REUSE (cc): reuse articulatory gestures common to the local (continuant consonant) system
>
> REUSE (all): reuse articulatory gestures common to the global (whole language) system.

The dual REUSE constraints imply that a segment must share gestural components with both the continuant subsystem and the larger system comprising continuants, occlusives, and vowels.

The articulatory ease constraint *HOLD is specified in order to provide for ranking of gestures. Only two relative rankings are necessary, this to distinguish between periodic articulations—which require positively more time in which a cycle of open-close is produced—and non-periodic articulations. These rankings are defined as follows:

> *HOLD (1): an articulator is not to be held in a non-neutral position for a positive length of time 1
>
> *HOLD (2): an articulator is not to be held in a non-neutral position for a positive length of time 2.

The universal ranking implication of *HOLD as defined is as follows:

*HOLD (2) >> *HOLD (1): it is worse to hold an articulator in a non-neutral position for a positive length of time 2 than for a positive length of time 1, where 2 is greater than 1.

Given these definitions and the gestural specifications of Table 4.1, it is possible to generate OT compatible models of how each language's continuant consonant inventory responds to the above constraints. Table 4.3 provides four articulatory grammars, one for each of the languages studied in this dissertation. Note that the coordination of all constraints is implied, save for that of * HOLD, as the purpose of OT presentation is not the generation of an optimal output, but the description of how each inventory responds to constraints and of the relations of segments within this system.

	*HOLD (2)	*HOLD (1)	REUSE(cc)	REUSE(all)
ENG {ɹ, l, s/z, ʃ, f/v, ð/θ}		******	****	**
FRE {ʁ, l, s/z, ʒ, f/v}		*****	*****	*
AMS {R, l, s/z, ʃ, f/v, x}	*	******	*****	***
BRAB {ʁ, l, s/z, ʃ, f/v, x}		******	*****	**

Table 4.3. Articulatory grammar of English (ENG), French (FREN), Amsterdam Dutch (AMS), and Brabant Dutch (BRAB)

Within the relational context of each system, there is one violation of *HOLD (1) per inventory member: each segment implies movement from a non-neutral position. Only Amsterdam Dutch contains a violation of *HOLD (2): this is the result of the periodic articulation of /R/.

Each language's continuant consonant inventory results in violation of REUSE, although some contain more violations than do others. Within the context of the continuant consonant inventory, English violates REUSE(cc) four times: /f, v/ and /ɹ/ share no articulator among others in the continuant system; /s, z/ and /ɹ/ share no articulator shape within the system. There are no violations of REUSE(cc) as the result of non-shared articulatory targets. Given the relations between the continuant consonants and the larger system of all articulatory segments in English, this inventory results in only two violations of REUSE(all), the articulator shape violations of /s, z/ and /ɹ/ mentioned above. French results in six violations of REUSE(cc): /z/ results in one for a unique articulator shape and /ʒ/ for its non-shared articulatory target; both /ʁ/ and /v/ violate this constraint for target and shape. Only the /z/ shape gesture results in a violation of REUSE(all). Amsterdam Dutch consists of five violations of REUSE(cc): /R/ articulator target (uvula) and shape, /s, z/ articulator shape, and /f, v/ articulator shape and target. /R/ target and shape, as well as /s, z/ shape violate REUSE(all). Brabant Dutch yields five REUSE(cc) violations: /ʁ/ target (velum-uvula), /s, z/ tongue shape, /ʃ/ target, and /f, v/ articulator shape and target. The first two of these are also REUSE(cc) violations.

Constraint interaction is also useful in presenting the articulatory dynamic of rhotics, in particular. Since these segments are shown to result in greater complexity and in a more multiplex articulatory mapping, the number of violations arising from these segments' relation to the larger system is of particular interest. Table 4.4 presents the articulatory grammars of the four rhotics in question.

	*HOLD (2)	*HOLD (1)	REUSE(cc)	REUSE(all)
English /ɹ/		*	**	*
French /ʁ/		*	**	
Amsterdam /R/	*	*	***	*
Brabant /ʁ/		*	*	

Table 4.4. Articulatory grammar of rhotics

The rhotic articulatory grammars paint a picture of these segments' place within their respective systems. Of particular concern are the REUSE constraints. The English /ɹ/ violates REUSE(cc) twice—a result of articulator and articulator shape—but REUSE(all) only once—due to the articulator shape. Although the tongue body is not common to any other continuant consonants, it is significant in the overall English articulatory grammar, used for /j/, /k, g/, and for vowels. The French /ʁ/ results in two violations of REUSE(cc), as both the articulator and target of articulation are not shared by other continuant consonants. However, this rhotic does not violate REUSE(all): the tongue body is used for /k, g/, for example, and these consonants share the velar/uvular place of articulation.

Amsterdam Dutch /R/ violates *HOLD(2), as mentioned above. This rhotic results in two violations of REUSE(cc)—for the uvular articulatory target and the bunched articulator shape; these are also violations of REUSE(all). Brabant Dutch presents a different case. Here, the rhotic /ʁ/ is violates only REUSE(cc), as its articulatory target is not shared by other continuant consonants. REUSE(all) is satisfied, however (for example by the systemic presence of /k/). In this form of Dutch, the rhotic displays a much different relation to its larger system and none of the negative relations that would set it apart as a peripheral gestural unit.

4.4.2. PERCEPTUAL GRAMMAR

The perceptual drive is not dependent upon gestures but upon the acoustic nature of segments and their interaction within the context of the continuant consonant system. 2.4.2 formulated a theoretical perceptual output using two constraints, MAINTAIN and *MINDIST, each of which was ranked using a hypothetical number of oppositions. As above, description of the perceptual dynamic of languages implies further specification of each of the constraints. However, it is also necessary to provide a means of ranking quantitative distance (for *MINDIST) and qualitative contrast (for MAINTAIN).

Determination of distance is accomplished by looking to each segment's acoustic profile (see Table 4.2) and noting the quantitative differences between a given consonant's formants and all corresponding formants within the system. For the purposes of the perceptual output grammar, a quantitative distance is defined as follows:

Distance (d) = 1, $d(A) \times d(B)$ *iff* $d \neq a(A) \wedge a(B)$

A quantitative distance (d) of one is defined as a positive distance (in Hz) in the union of systemic sets (A) and (B), if and only if the component (a) of set (A) is not present in (B).

Given this definition, a quantitative distance is understood to be any measurement of quantity in a given systemically related segment that is positively distinct with regard to all quantity measurements in another systemically related segment.

Tables 4.5 through 4.8 provide the quantitative distances resulting from segment interaction in each language system.

Formalization of quantitative distances reflects the perceptual mapping of 4.2.1. Considered in terms of the frequency of maximal acoustic energy, the interaction of some segments results in no distance distinction: this is the case for /s, z/ and /f, v/ in all systems. Other segments, such as the English /ɹ/ and /l/, are minimally distant. Still more, such as /l/ and /f, v/, present the maximal number of distances under consideration.

	ɹ	**l**	**s, z**	**ʃ**	**f, v**	**ð, θ**
ɹ	--	1	3	3	3	3
l	1	--	3	2	3	2
s, z	3	3	--	2	0	1
ʃ	3	2	2	--	2	0
f, v	3	3	0	2	--	1
ð, θ	3	2	1	0	1	--

Table 4.5. Quantitative distances, English

	ʁ	**l**	**z**	**ʒ**	**v**
ʁ	--	1	3	2	3
l	1	--	3	2	3
z	3	3	--	2	0
ʒ	2	2	2	--	2
v	3	3	0	2	--

Table 4.6. Quantitative distances, French

	R	l	s, z	ʃ	f, v	x
R	--	0	3	2	3	1
l	0	--	3	2	3	1
s, z	3	3	--	2	0	1
ʃ	2	2	2	--	2	2
f, v	3	3	0	2	--	1
x	1	1	1	2	1	--

Table 4.7. Quantitative distances, Amsterdam Dutch

	ʁ	l	s, z	ʃ	f, v	x
ʁ	--	1	3	2	3	1
l	1	--	3	3	3	1
s, z	3	3	--	2	0	1
ʃ	2	3	2	--	2	1
f, v	3	3	0	2	--	3
x	1	1	1	1	3	--

Table 4.8. Quantitative distances, Brabant Dutch

Whereas quantitative distance looks only to the measurement of formant peaks and disregards the acoustic nature of the formants, qualitative contrast looks to both. Quality is an important factor in perception, as listeners repair the type of sound, as well as its frequencies, to categorized patterns. For the purposes of this dissertation, a qualitative contrast is defined as follows:

Contrast (c) = 1, $c(A) \times c(B)$, *iff* $c \neq a(A) \wedge a(B)$

A qualitative contrast (c) of one is defined as a positive contrast in the union of systemic sets (A) and (B), if and only if the qualitative component (a) of set (A) is not present in (B).

A contrast is thus a qualitatively distinct, quantitative component of one systemically related segment with regard to another. Qualitative contrasts permit the differentiation of quantitatively similar—or even identical—segments, such as /s, z/ and /f, v/. Tables 4.9 through 4.12 present tabulations of the qualitative contrasts within the four subject systems.

	ɹ	**l**	**s, z**	**ʃ**	**f, v**	**ð, θ**
ɹ	--	1	3	3	3	3
l	1	--	3	3	3	3
s, z	3	3	--	2	3	3
ʃ	3	3	2	--	3	3
f, v	3	3	3	3	--	1
ð, θ	3	3	3	3	1	--

Table 4.9. Qualitative contrasts, English

	ʁ	**l**	**z**	**ʒ**	**v**
ʁ	--	3	3	3	3
l	3	--	3	3	3
z	3	3	--	2	3
ʒ	3	3	2	--	3
v	3	3	3	3	--

Table 4.10. Qualitative contrasts, French

	R	**l**	**s, z**	**ʃ**	**f, v**	**x**
R	--	0	3	3	3	3
l	0	--	3	3	3	3
s, z	3	3	--	2	3	2
ʃ	3	3	2	--	3	2
f, v	3	3	3	3	--	3
x	3	3	2	2	3	--

Table 4.11. Qualitative contrasts, Amsterdam Dutch

	ʁ	l	s, z	ʃ	f, v	x
ʁ	--	3	3	2	3	1
l	3	--	3	3	3	3
s, z	3	3	--	2	3	1
ʃ	2	3	2	--	3	1
f, v	3	3	3	3	--	3
x	1	3	1	1	3	--

Table 4.12. Qualitative contrasts, Brabant Dutch

In order to integrate systemic quantity and quality observations into a descriptive framework and provide perceptual grammars for each language, it is necessary to more precisely define each of the relevant constraints. These definitions take into account the quantitative distances and qualitative distinctions or contrasts and provide for a universal ranking of each. The first constraint looks to the acoustic profile of segments and states that there should be no quantitative distance between related segments. *MINDIST provides that a segment should not be distant from any other segment, reflecting the need for system focus. This constraint is reformulated as follows:

> *MINDIST (1): do not allow one minimum quantitative distance between a segment and all other segments
>
> *MINDIST (2): do not allow two minimum quantitative distances between a segment and all other segments
>
> *MINDIST (3): do not allow three minimum quantitative distances between a segment and all other segments.

A segment that has one or more minimum distances from another segment violates *MINDIST (1) one time. Likewise, a segment that has two minimum distances between itself and all others yields one violation each of *MINDIST (2) and (1). One segment may result in numerous violations. For instance, the English /ɹ/ results in three violations of *MINDIST(3).

Formalization of the perceptual drive in 2.4.2 states that, assuming perceptual distances are present in a system, the best functional system also avoids the superfluous use of these. Given that it is preferable to maintain as few quantitative distances within a system as possible—thus making more efficient use of perceptual space—the following universal ranking hierarchy is established:

> *MINDIST (3) >> *MINDIST (2) >> *MINDIST (1): it is less bad to have one quantitative distance between a segment and all others than to have two; it is less bad to have two quantitative distances between a segment and all others than to have three.

Such ranking allows for differentiation between segments without the superfluous use of quantitative distances, thereby promoting an efficient use of acoustic space. Note that no further specification of *MINDIST is used in this study, as no segment's interaction presents more than three quantitative distances.

MAINTAIN is also specified according to relational differences that are observed in the output form of each language. Here, reference is made to the maintenance of a minimum number of qualitative contrasts between a segment and all other related segments and is provided as follows:

> MAINTAIN (1): maintain at least one qualitative contrast between a segment and all other segments
>
> MAINTAIN (2): maintain at least two qualitative contrasts between a segment and all other segments
>
> MAINTAIN (3): maintain at least three qualitative contrasts between a segment and all other segments.

Given this definition, any segment that does not maintain a minimum of one qualitative contrast between itself and all others of its system will violate MAINTAIN (1). Likewise, a segment that does not maintain at least two contrasts violates MAINTAIN (2), and so on.

Because a functional system prefers the maintenance of a maximal number of qualitative contrasts, the universal ranking of MAINTAIN constraints is different with regard to those of *MINDIST.

> MAINTAIN (1) >> MAINTAIN (2) >> MAINTAIN (3): it is better to maintain at least three qualitative contrasts between a segment and all others than it is to maintain two; it is better to maintain at least two qualitative contrasts between a segment and all others than it is to maintain one.

Such ranking promotes the use of the greatest number of qualitative contrasts and promotes the relational distinction of segments.

The continuant consonant inventory of each subject is described using the interaction of *MINDIST and MAINTAIN, the result of which is presented in Table 4.13. This tableau describes the grammar of each perceptual system and does not follow the generative pattern typical of OT analyses. For this reason, no constraint coordination is provided, nor are *MINDIST and MAINTAIN ranked with respect to each other.

	Ma (1)	Ma (2)	Ma (3)	*MD (3)	*MD (2)	*MD (1)
Eng. {ɹ, l, s/z, ʃ, f/v, ð/θ}		****	*(6)	*(12)	*(21)	*(26)
French {ʁ, l, z, ʒ, v}			**	*(8)	*(16)	*(18)
Amst. D. {R, l, s, ʃ, f, x}	**	**	*(8)	*(8)	*(18)	*(26)
Brabant D. {ʁ, l, s, ʃ, f, x}		*(6)	*(10)	*(12)	*(18)	*(28)

Table 4.13. Perceptual grammar of American English, French, Amsterdam Dutch, and Brabant Dutch continuant consonant inventories. *MINDIST is abbreviated as *MD, MAINTAIN as Ma. Violations greater than five are represented by * and a number in parentheses, indicating the total violations of the relevant constraint.

Interaction of *MINDIST and MAINTAIN constraints describes the relative efficiency of each of the perceptual systems, when viewed as a system of relations. Even in systems with the same number of participant members, such as English and the two forms of Dutch, differences are noted in the perceptual grammar. For instance, English does a better job at providing for qualitative contrast between segments, as seen in the number of MAINTAIN violations, whereas Amsterdam Dutch does the worst. All three are similar in terms of their use of quantitative distance, but the relational system of Brabant Dutch results in the greatest number of superfluous distances, as seen in the number of *MINDIST(3) violations.

Similar to the presentation of rhotic articulatory grammar in 4.4.1, it is possible to look at the particular perceptual relations between these segments and the systems in which they participate. Table 4.14 presents the perceptual grammar of rhotics.

	Ma(1)	Ma(2)	Ma(3)	*MD(3)	*MD(2)	*MD(1)
English ɹ		*	*	****	****	*****
French ʁ				**	***	****
Amsterdam Dutch R	*	*	*	**	***	****
Brabant Dutch ʁ		*	**	**	****	*****

Table 4.14. Perceptual grammar of rhotics. *MINDIST is abbreviated as *MD, MAINTAIN as Ma.

A perceptual grammar of rhotics defined according to the violation or satisfaction of *MINDIST and MAINTAIN constraints captures the relational dynamic of these segments. While all rhotics in this study result in several *MINDIST(3) violations, the number of higher-ranking *MINDIST(1) and (2) violations is different and it would seem that English does the worst job of avoiding superfluous use of quantitative distance. Amsterdam Dutch, on the other hand, presents the only violation of MAINTAIN(1), this due to the qualitative confusion of /R/ and /l/.

Chapter Five uses the foundational grammars presented above and in 4.4.1, demonstrating how relational grammars can account for instances of variation seen in rhotic behavior.

4.5. Relational integrity: towards a relationally-defined segment

This chapter advances the notion that a segment is defined by its relation—both positive and negative—to other segments of its system and not by any internal structure, be it gestural or acoustic in nature. I give this notion the title of relational integrity, stating that segments need not be defined with reference to universal or pre-established categories or features. I assert that segmental integrity—i.e. the definition of what is or is not a given segment within the context of a larger system—derives from the organic linguistic system and from the interaction of its constituent members. Relational integrity thus relies on both positive and negative definition.

It is useful to contrast the theoretical propositions of relational integrity and of the relational segment with antecedent definitions of the phonological segment, both those that have inspired it and those opposing it. Chapter Two provides an overview of several functionalist (or proto-functionalist) approaches to language. The most important of these is Passy (1891). It is fair to say that my theory of relational integrity is directly related to his declaration that language consists of a systematization of contrasts: "*dans la langue, il n'y a que des différences*." Other works also contribute to relational integrity as both an innovative and ancient idea.

Trubetzkoy (1939) refers to the phonological system as one of opposition and correlation. For instance, he provides for analysis of the German consonant system according to two types of contrast—one- and multidimensional oppositions. These are bound together or rendered into a systemic whole via correlation of oppositions one to another. While Trubetzkoy's proposition of forty one-dimensional and 150 multidimensional oppositions appears overly complex, given the limited nature of human consonant inventories and the need for relative simplicity in systemic relations, it is nonetheless a brilliant work whose influence is still considerable in our time. Indeed, his understanding of the interaction of both productive and auditory properties of sounds serves as a critical backdrop for functionalist and non-structuralist theories such as Martinet and Boersma, as well as the present dissertation.

More recent (and traditionally accepted) descriptions of segments and inventories have relied on the use of ad hoc features. While none assume all features are used by all languages or within all systems, the definition of features is bound up in the local application of universally posited structures. Jakobson, Fant and Halle (1967) make use of no fewer than 53 features derived almost exclusively from the acoustic and perceptual properties of segments. This contrasts to Chomsky and Halle (1968), who provide forty features, largely derived from gestural characteristics and articulatory behavior. Later works integrate a mix of features into geometric models, providing a visual representation of the internal structure of phonological units. Clemens (1985) was the first to define a segment using three-dimensional models. His model of feature geometry derives from traditional structuralist approaches, namely Chomsky and Halle, and has proved useful in describing the phonotactic interaction of segments.

The proposition and use of externally motivated features—whether perceptual or articulatory in nature—fall prey to an inherent flaw. All such models begin with the assumption that features are universal and that they must be an implicit—if unrealized—component of the structure of all languages, of all inventories, and of all members of these inventories. The feature is, in these theories, not a structure that emerges from language but one that exists above it. Whether or not a particular language or dialect makes us of the feature or ignores it is irrelevant. This highlights a theoretical flaw of primarily structural approaches. Segments are defined via the use of synthetic features not innate to the organic environment to which they are supposed to belong, i.e. the system within which they exist. As a result, it is continually necessary to add and subtract primary or secondary features in order to enlarge the scope of feature theory and to explicate language data. Even models that provide for underspecification—radical or otherwise—fail on this account.

The functionalist approach outlined in this dissertation differs from traditional structural approaches, inasmuch as it does not admit the universality of such commonly held structures as features and denies the existence—primal or otherwise—of nodes. My approach is not the first to call into question synthetic application of features. Boersma opposes the physical realities that define articulation and perception, contrasting this to the idiosyncrasies of language data and denying traditional understanding of features.

> **The functional view: no universal phonological feature values**
>
> The continuous articulatory and perceptual phonetic spaces are universal, and so are the constraints that are defined on them; the discrete phonological feature values, however, are language-specific, and follow from the selective constraint lowering that is characteristic of the acquisition of coordination and categorization. (1998: 172, bold of the original)

He also concludes that the nodes of feature geometry, specifically the place node as seen in the Clemens and Hume model, do not have any basis in the reality of speech (1998: 23-24, 441-42, 463). Rather than rely on hybrid nodes, which mix gestural and perceptual acts (ergo, speaker and listener), Boersma provides separate grammars for each linguistic component of communication, much like the articulatory and perceptual grammars of 4.4.

Relational integrity as I propose it within the context of a functionalist understanding of language avoids these difficulties by refusing the universal availability of features or, more radically, by denying the traditional notion of feature altogether. I assume that features do not exist outside the context of a system of opposition and correlation, in Trubetzkoy's terms, or a system of differences, to borrow from Passy. Thus, it is not only unnecessary to underspecify the English continuant system for the articulatory feature [periodicity], for instance; it is entirely incorrect to associate this feature to this system at all. English is not underspecified, but aspecified, with regard to periodicity.

For clarity, all further mention of the categorized components or features of a presumed segment—both articulatory and perceptual—should be understood as follows:

> **Feature** (relational definition): A feature is a gestural or acoustic component of a relationally distinct sound that emerges from that segment's inclusion in a larger, organic linguistic construct and from the discriminability of that segment within the system. E.g. the gestural component "tongue bunched contour" emerges as a feature of English by virtue of /ɹ/; the acoustic component "2790-3200 Hz" emerges as a feature of Brabant Dutch by virtue of /ʁ/.

Following this definition, questions of feature universality are removed from the phonological equation. The articulatory grammar of the French, for example, provides for the relation of segments according to three target features, three articulator features, and two articulator contour features. That these parameters are similar to those of English, Dutch or any other language does not imply that features are universal. Rather, it is quite natural that languages in such close historic, social, and genetic contact should display similarities in their parametric, relational organization.

The traditional understanding of a phonological segment is also called into question by the relational approach and by its redefinition of features, the usual building blocks of grammars. It has been stated at several points throughout this dissertation that there is no inherent structure to segments, but that the definition of what is or is not a given segment is actually inherited from a larger, organic system of relations. The following provides a succinct relational definition of the segment:

> **Segment** (relational definition): a participant member of an organic system whose distinction results from the ensemble of similarities and differences—both perceptual and articulatory—vis-à-vis all other members of that inventory. (N.B. segments are abbreviated with IPA characters by convention.)

Defined relationally, a segment emerges from its interaction within the system.

The example of French is sufficient to underscore the importance of systemic definition and of phonetic grounding to the theory of relational integrity. Because a segment may be defined only in relation to other segments shared by a common system, each segment must be mapped before any definition of any one particular member of the system can be advanced. Such phonological grounding does not deny that languages share similar articulatory and perceptual characteristics, but refuses to allow for the definition of the component of one organic whole by virtue of the component of another. Stated in more simplistic terms, a segment and the systemic organization of French are wholly unconcerned with a particular segment in or the systemic organization of English, Polish, Malay or any other language. French is organized according to the articulatory and perceptual functions attributed to it by its speakers.[19]

Relational definitions of the phonological feature and segment lead naturally to the question of how such locally defined entities can be related across languages. An example of this type of investigation is seen in a comparison of similarities between sibilant fricatives /s, z/, for instance those of English, French and Dutch. If each such segment is defined according to the similarities and differences correlated to other segments, a relational approach can compare these dynamics and posit a relational pattern. This exemplifies the notion of relational integrity, which I define as the following:

> **Relational integrity**: The congruence of similarities and distinctions presented by two or more segments, each defined according to organic relations to a singular system, across two or more such systems. E.g. the relational integrity of /v/ in English, Dutch, and French derives from the definition of the English /v/, the Dutch /v/, and the French /v/ and the mutual congruence of these relational definitions.

Whereas the definition of a particular segment and its integrity within a systemic context is horizontal—looking only at the tacit relationship of similarities and differences within the organic whole of the system—the comparison of such relationships with similar segments from different languages is entirely vertical.

[19] Taken to its fullest extent, a theory of relational integrity calls into question formal notions of language, dialect and sociolect, reconciling definitions of the linguistic code with those provided by sociolinguistics. A linguistic code understood in these terms is not an absolute, but is borne from the interaction of speakers sharing or deferring to a commonly shared norm.

Relational integrity provides for a relational class, the delimitation of a type of sound based on how these sounds interact across linguistic systems.

It is the overall system of relational dynamics—both articulatory and perceptual—that provides for the vertical relation of segments and for the relational integrity of segments from different systems. This notion, that similar segments constitute a relational—as opposed to static—category opens the door for a greater phonological understanding of segments and their behavioral characteristics.

4.5.1. REDEFINING RHOTICS: VERTICAL RELATIONAL INTEGRITY

There have been but two viable attempts at defining rhotics as a class of sounds, distinct from all others, those of Lindau (1985) and Walsh Dickey (1997). While description of these is provided in Chapter One, I believe it useful to briefly return to these studies and contrast their understanding of what is and is not a rhotic to my own, relationally biased approach.

The most prominent contrast between my approach to rhotic classhood and more traditional ones is seen in Walsh Dickey (1997). Her definition of the feature [rhotic] relies entirely on cross-linguistic observations of rhotic behavior and on the position of multiple features—most importantly a secondary laminal node—onto the geometric structure of rhotics. Despite her faultless application of structural theory and a logically presented case for this feature, it is impossible to reconcile our two works. This is due to the functionalist foundation of this dissertation and to my own assertion that features do not exist outside the context of an organic, linguistic whole. Where Walsh Dickey looks for evidence for rhotic classhood in cross-linguistic similarity of posited structure, relational integrity denies any possible comparison.

Lindau (1985) offers a theory of rhotics more reconcilable to the one outlined herein. She provides for the unity of diverse segments by establishing a series of parametric relations, i.e. multiplex similarities between different rhotics. My own understanding of relational dynamics and theory of relational integrity owes much to her work, but differs substantially from it in two regards. Firstly, the bases for Lindau's relations do not derive from organic systems but from synthetic systems, i.e. from observations of similarities seen in the grouping of rhotics, absent from their larger systemic context. She provides for the parametric relation of the French [ʁ] with the Swedish [ʀ], for instance, without reference to the larger French and Swedish systems. Relational integrity as I have defined it in this chapter cannot depart from this point. A secondary difference between my approach and that of Lindau is seen in the combination of articulatory and perceptual characteristics as unique parameters. Rather than allow for the confusion of two separate biomechanical systems—one depending on gestures and the other on acoustics—I provide for their rigorous distinction.

Given a relational definition of the segment as a construct of differences and similarities within the context of an organic system, no structural explanation can be put forth to account for the similarities noted between rhotics in different languages. One may, however, look to the pattern of differences and similarities present within several systems to provide a clue as to the relational "place" within a language system and thus to the relational pattern of similar, potentially rhotic segments in one or more other language systems.

The articulatory drive provides a first clue as to the relational integrity of rhotics. Considered in terms of the articulatory drive and grammar, it is noted that no rhotic is defined horizontally by a singular articulatory component. These segments are, in fact, the only ones in question that are not named for a singular gestural trait (e.g., "dental," "grooved," or "lateral"). Rather than derive from a specific component, a pattern of vertical component distinctions characterizes the rhotics. Definitions of the articulatory integrity of each rhotic are given as follows:

> **English /ɹ/**: "≠ /s, z/ (articulator, target, shape), /ʃ/ (articulator, shape), /f, v/ (articulator, target, shape), /ð, θ/ (articulator, target, shape), /l/ (articulator, target, shape); = /ʃ/ (target)"
>
> **French /ʁ/**: "≠ /z/ (articulator, target, shape), /ʒ/ (articulator, target), /f, v/ (articulator, target), /l/ (articulator, target); = /ʒ/ (shape), /l/ (shape)."
>
> **Amsterdam Dutch /R/**: "≠ /s, z/ (articulator, target, shape), /ʃ/ (articulator, target, shape), /f, v/ (articulator, target), /x/ (shape), /l/ (articulator, target, shape); = /ʃ/ (target, shape), /x/ (articulator, target, shape), /l/ (shape)."
>
> **Brabant Dutch /ʁ/**: "≠ /s, z/ (articulator, target, shape), /ʃ/ (articulator, target), /f, v/ (articulator, target), /x/ (target), /l/ (articulator, target); = /ʃ/ (shape), /x/ (articulator, target, shape), /l/ (shape)."

These system specific integrities reveal congruence with regard to rhotics. Within the context of the four systems above, the following articulatory relational integrity of rhotics is provided:

Rhotic (articulatory): "≠ /s, z/ (articulator, target, shape), /ʃ, ʒ/ (articulator), /f, v/ (articulator, target), /l/ (articulator, target)."

In terms of articulatory integrity, the rhotics of this study constitute a relational class by virtue of articulator distinction in relation to {s/z, ʃ, f/v, l}, target distinction in relation to {s/z, f/v, l}, and shape distinction in relation to {s/z}. They present no positive vertical relations. The articulatory rhotic might therefore be described as "not /s, z/, /ʃ, ʒ/, /f, v/, or /l/."

Definition of perceptual relational integrity is similar to that of articulatory integrity; given the diversity of perceptual profiles associated with the rhotics of this study, no singular pattern of perceptual relation emerges from these segments. Perceptual integrity derives from positive and negative relations (i.e. systemic similarities and distinctions). The following are formal definitions of specific rhotic integrity:

English /ɹ/: "≠ /s, z/ (3 quantities, 3 qualities), /ʃ/ (3 quantities, 3 qualities), /f, v/ (3 quantities, 3 qualities), /ð, θ/ (3 quantities, 3 qualities), /l/ (1 quantity, 1 quality); = /l/ (2 quantities, 2 qualities)."

French /ʁ/: "≠ /z/ (3 quantities, 3 qualities), /ʒ/ (2 quantities, 3 qualities), /f, v/ (3 quantities, 3 qualities), /l/ (1 quantity, 3 qualities); = /ʒ/ (1 quantity), /l/ (2 qualities)."

Amsterdam Dutch /R/: "≠ /s, z/ (3 quantities, 3 qualities), /ʃ/ (2 quantities, 3 qualities), /f, v/ (3 quantities, 3 qualities), /x/ (1 quantity, 3 qualities); = /ʃ/ (1 quantity), /x/ (2 quantities), /l/ (3 quantities, 3 qualities)."

Brabant Dutch /ʁ/: "≠ /s, z/ (3 quantities, 3 qualities), /ʃ/ (2 quantities, 2 qualities), /f, v/ (3 quantities, 3 qualities), /x/ (1 quantity, 1 quality), /l/ (1 quantity, 3 qualities); = /ʃ/ (1 quantity, 1 quality), /x/ (2 quantities, 2 qualities), /l/ (2 quantities)."

The vertical congruence of these system specific integrities results in the following perceptual relational integrity of the rhotics in question:

> **Rhotic (perceptual)**: "≠ /s, z/ (3 quantities, 3 qualities), /ʃ, ʒ/ (2 quantities, 2 qualities), /f, v/ (3, quantities, 3 qualities), /l/ (1 quantity, 1 quality); = /l/ (2 quantities)."

In contrast to the articulatory relational integrity of the rhotics {ɹ, ʁ, ʀ}, perceptual integrity presents one positive vertical relation, that of rhotics to /l/ by virtue of cross-linguistically shared quantitative distances. This is not, however, the most important component of perceptual integrity; rhotics are in all four systems maximally contrasted to /s, z/ and /f, v/ and rather distinct from /ʃ, ʒ/. These negative relations are as much a part of rhotic integrity as are any positive ones.

Relational integrity provides for a rather simple answer to the question of apparent rhotic classhood posed in Chapter One. The similarity of one "r-like sound" to another is not due to either sound's internal structure, but due to these segments having congruent relational integrity within larger systems. Stated rather simplistically, the French /ʁ/ is similar to the English /ɹ/ because these segments are respectively correlated to French /l/, /z/, /v/, etc… and to English /l/, /z/, /v/, etc. Another simplified definition of the French and English rhotics provides that each is different from its related /l/, /z/, /v/, etc. and it is this differentiation which lies at the base of their system-external similarity.

While it would seem that such a series of cumbersome, yet intuitively straightforward definitions has little physical ground on which to stand, the results of articulatory and perceptual mapping and constraint based, grammatical output indicate that these intuitions are theoretically sound. Each of the concerned rhotics occupies a particular place in its system's articulatory and perceptual grammar. The remarkable similarity of the rhotic relational dynamic—whether peripheral or central—is the only means of explaining similarities observed across languages available to a functionalist.

While relational integrity provides a means to explain cross-linguistic similarities, it does not posit any universal feature [rhotic]. This is an innovative understanding of sound segments, in general, and of rhotics, in particular. In the following chapter, I demonstrate how a relational understanding of rhotics can be used to account for variations in rhotic behavior. From such demonstrations, I advance a theory of relational integrity that, when properly grounded in functionalism and in phonetic observations of linguistic output, can and does allow for the definition of what is and is not a rhotic.

CHAPTER FIVE. THE IMPLICATIONS OF RELATIONAL PHONOLOGY: ACCOUNTING FOR RHOTIC VARIATION

5.1. Introduction

Chapter Four articulates working definitions of relational phonology and relational integrity, providing articulatory and perceptual mappings and defining the articulatory and perceptual grammars of English, French, and Dutch with violable, systemic constraints. Within these phonologies, I have shown that /r/ constitutes a relational class, i.e. a type of sound segment whose relation to other, homosystemic members is predictable and congruent in different linguistic systems. This chapter asks a different question of rhotics and of the relational theory. Namely, if one adheres to the understanding of language as being a dynamic system of differences and similarities and of sound segments as being defined solely thereby, how are cases of variation to be described and explained?

Traditional phonological analyses of variation depend on a defined input or underlying form and apply constraints or ordered rules to generate variant outputs or surface forms. As mentioned previously, these phonologies define both input and output segments with an amalgamation of articulatory and perceptual features, assumed to be universally present and/or available, regardless of the linguistic system in question. Chapters Two and Four deny feature universality, assuming that these are emergent properties of the linguistic system and its competing drives.

Relational phonology, as I propose it, presents a radically different approach to variation. Segments are defined within an organic context in terms of their relational integrity, i.e. those articulatory or perceptual, emergent features that allow for their distinction within the system. This definition captures within its boundaries a considerable amount of variation that, from a phonetic point of view, has not been and will not be treated in this work. For example, the French /z/ presents a great deal of quantitative perceptual variation dependent on tautosyllabic vowels. A relational approach to the definition of /z/ in this system does not concern itself with this, insomuch as this variation is of no consequence for the distinction of the segment vis-à-vis /ʒ/, /v/, etc. The description, evaluation, and explanation of such variation constitute a phonetician's work and have only passing interest to the present study.

The task of a relational phonology is therefore to describe and explain those instances of variation that result in grammatical changes, i.e. differential satisfaction and/or violation of those constraints applicable to the articulatory and perceptual grammars of the system in question. A relational approach to language also asks what systemic mechanisms allow for or militate against variation in the output form of relationally defined input segments.

Three different types of rhotic variation are seen in the languages studied in this dissertation. In Amsterdam Dutch, a number of questions have been left unresolved from previous chapters, especially with regard to the variable articulatory nature of this language's rhotic and to the potential perceptual confusion of /R/ with other segments, namely /x/ and /l/. American English presents a different type of variability, where standard forms of the language use /ɹ/ in all instances, but where several dialects drop or reduce rhotics in word-final environments. From these examples of allophony, a relational model is provided to explain how such variation is allowed by the articulatory and perceptual grammars and how variation responds to the competing systemic needs of efficiency and discriminability.

A similar instance of variation is addressed in regard to the French /ʁ/, described as being either a fricative or approximant and subject to a great degree of alternation that, when viewed from a relational perspective, has no phonological motivation. This variation is subject to no clear phonotactic or environmental rules: the choice of approximant verses fricative appears to be a case of free variation, at least within the context of a phonological analysis. I show how a functionalist understanding of language and relational approach to the sound segment allows for a better understanding of so-called free variation. Furthermore, I clearly state how such apparent failure on the part of a relational phonological analysis does not detract from the validity of the approach I evolve in this dissertation.

5.2. Accounting for relationally defined variation within a constraint based grammar

Optimality Theory has gained wide acceptance among phonologists as a means of expressing naturalness in rule-based, generative models of linguistic behavior. In nearly all OT models, constraints belong to one of two classes:

faithfulness constraints, which promote output candidate congruence to a specified input; and markedness constraints, which promote variation in the output candidate set as a means of satisfying extra-linguistic requirements place on the communicative system. Each type of constraint may be stated positively (“do constraints”) and negatively (“do not constraints”), the latter abbreviated with an asterisk (*). All constraints are violable and, at some level of theoretical interpretation, universal.

Analyses in this chapter represent a departure from conventional OT in two regards. Firstly, there is a limited role for faithfulness constraints within a relational approach, owing to the diversity of input variation captured by articulatory and perceptual definitions of phoneme segments. The role of FAITH constraints is discussed further in 5.2.1, below. Also, as in Chapter Four, I make use of only those constraints that can be motivated by supralinguistic principles, deriving from the biomechanical processes of articulation and perception. As such, it is perhaps best to refer to the approach used in this chapter as being inspired by, rather than emulating classical OT.

The functionalist principles and language specific implications presented in 2.3 and 2.4 assert that constraints placed upon linguistic function, hence on linguistic form, are universal and must be motivated, rather than adduced. For the purposes of this chapter, it is necessary to enlarge the scope of investigation and capture a number of linguistically external, universal principles of human biomechanical activity. Quite naturally, these are applied to linguistic activity.

Discussion of Lindblom's (1990) Hyper-Hypo (H&H) is introduced in Chapter Two and contributes to the articulation of functionalist principles and implications. Although H&H predated the OT model, these are not mutually exclusive. An OT approach to language form assumes competition between universally existent constraints; ideally, the motivation for these constraints would come not from particular linguistic forms but from supralinguistic principles. H&H offers a succinct formalization of one such principle, assuming that speech is accomplished along a continuum of hyper (output-oriented) and hypo (system-oriented) speech and that the selection of speech modes is subject to biomechanical principles.

Lindblom (1990) presents a phonetic, articulatory perspective, considering relative differences in the amount of articulatory effort implied by a gesture and the perceptual effect of more or less effort in a given speech act. In the former case, the speech signal is of primary importance. Relatively greater motor control and articulatory effort are put into play in order to produce an acoustic signal that can be more easily discriminated. Hyper speech is purpose driven: how an utterance is produced is just as important as what is said, if not more so. In the latter instance, the speech signal is only important inasmuch as it provides sufficient discriminability. Given that lexical processes interfere with and assist in perception, the output orientation is allowed to suffer. Hypo-speech is driven by the need for gestural economy and reflects a default state of speech production.

H&H theory enjoys the support of both phonetic and phonological literature. Lindblom (1990: 428-430) applies H&H to the case of phonetic vowel reduction in clear speech. Subjects produced vowels with less undershoot in hyper speech tokens, where they were asked to speak as clearly as possible. Citation forms, by contrast, showed more vowel formant undershoot. An earlier version of H&H, looking only to the economy of speech gestures, applies productive considerations to the cases of assimilation and segmental organization in consonant clusters, showing that degrees of gestural difficulty account for phonological variation (Lindblom 1983).

Questions of sound change are explained an H&H approach in Lindblom et al. (1995), reformulating the system and signal orientations as "what" and "how" modes, respectively. The authors distinguish three types of phonological processes applicable to diachronic change: prosodic processes that affect speech timing and syllable structure; weakening processes, such as assimilation, reduction, and lenition; and strengthening processes, such as vowel shift and consonant fortition (1995: 17-18). Each instance of change, regardless of type, is filtered by the what/how modes or system/signal orientations and both the listener and speaker are considered to be sources of sound change. A "what mode" favors weakening and prosodic processes that allow for more system efficiency, so long as the potential for confusion is minimal. On the contrary, fortition and strengthening processes are born from the "how mode," which seeks to avoid instances of confusion.

Kohler (1990) presents a generative approach to synchronic segmental reduction of /r/ in German connected speech that, while not explicitly making use of H&H, reflects this theory's stated need for the minimization of energy expenditure. He demonstrates that, in speech tokens where listeners allow (or do not forbid) lenition and elision processes, these will take place to the greatest extent possible as a means of achieving gestural economy. Restraint on economy derives from the context of the speech act and from phonological rules. Reduction processes exist on a continuum and economy is gradient, subject to goal-oriented conditions. His approach does not explicitly treat perceptual discriminability and states only that the goal of articulation is to provide minimal perceptual contrast in a listener-oriented mode.

Boersma (1998) approaches allophonic variation from a perceptual bias, providing for a learning step or stochastic evaluation of potential candidates in a functionally defined, generative output mechanism (332-338). He uses listener-oriented markedness constraints, abbreviated as *WARP, that militate against the perceptual categorization of maximally dispersed acoustic values. *WARP promotes efficiency in the perceptual system, eliminating those variants that are too far from a specified input and, for that reason, difficult to repair.

The economy of speech gestures and the discriminability of perceptual cues provide the basis of Boersma's optimization strategy (2000). Evaluating instances of sound change, he demonstrates that the quest for maximal ease of articulation and perception place the two drives of linguistic production at eternal odds. The articulatory system seeks to do as little as possible, avoiding expenditure to the greatest degree possible: the perceptual system seeks to maximize acoustic differences, using as many as possible, even if this implies more articulatory energy is to be used. Within this paradigm, optimization results from the best-case scenario of articulatory and perceptual outputs. Because optimization is never fully realized, a tug-of-war is created, motivating systemic innovations and, inevitably, variation. Innovations do not, however, respond perfectly to either articulatory or perceptual needs; optimization is therefore an unending process, continually at play in the evolution of language.

All of the above approaches to questions of variation—whether synchronic or diachronic—assert that variation is the result of competing needs to which the linguistic systems must respond. These needs are motivated by linguistically external principles of economy and emphasis that, when applied to the communicative act, constrain what a system is allowed to do with regard to the needs of speakers and listeners.

Both articulatory and perceptual systems will accomplish a given task using the least amount of energy possible. In circumstances requiring more precision or discriminability, higher cost activities are allowed. These considerations are provided in Chapter Two as Functionalist Principles 4 and 5 and as Implications 4 through 6. In order to apply H&H to an OT model of allophony, the tension between system and output orientations and economy and plasticity must be captured by unary constraints. For the cases at hand, these violations must distinguish between two phonotactic environments. I have chosen to refer to these as primary and secondary environments, although the more traditional terminology of "strong" and "weak" prosodic environments is not entirely without motivation.

For the purposes of this section, a primary phonotactic environment is considered to be one where the need for distinction and discriminability are relatively greater. Beckman (1998) asserts that the onset position enjoys a prosodically privileged position in the word. This is due to the perceptual importance of lexical information in word and syllable initial positions. She further provides evidence for codas (among others not relevant to the present study) as being non-privileged positions. In her analyses, which use a classical OT framework, segments in privileged positions must respond to higher ranked, prosodically specified Faithfulness constraints. Conversely, segments in non-privileged positions are subject to corresponding constraints of lower ranking. Steriade (1995) shows that, within a classical OT model, faithfulness constraints and perceptually oriented markedness constraints apply differently to onset and coda segments. Additional evidence for distinction between onset and coda environments is seen in Kingston (1990), who discusses over-articulation (corresponding to hyper or output orientations, above) in word-initial and onset environments.

In all discussion below, I refer to onset environments as primary and to coda environments as secondary. The reader will notice that there is only an approximant and certainly not absolute correspondence between these terms and those of Beckman or the more traditional “strong” and “weak” positional titles. In the case of word-initial and onset segments, little external lexical information is available to the listener and the system shows demonstrable output bias. This is considered a primary segmental environment. Intervocalic segments are also deemed primary.[20] In both of types of primary environments, the articulatory and perceptual drives provide for an output, hyper orientation. Depending on systemic needs and the range of possible, systemically defined variants, this orientation responds to the need for greater perceptual discriminability in environments that are otherwise impoverished of lexical information.

In the secondary environments, i.e. coda and word-final position, lexical information provides indirect evidence to listeners and speech is allowed to default to a system orientation. Because listeners may access a wealth of linguistic and contextual information to aid in perception and increase the likelihood of lexical repair, speakers are allowed to use less effort, i.e. less articulatory energy and resultant perceptual contrasts.

Considerations of primary versus secondary environments mirror the two poles of H&H, as well as the tension underlying the "what" and "how" modes of speech production and perception. I describe this competition with two series of constraints. These are articulatory and perceptual markedness constraints, indicating what a systemically defined segment should provide in a given particular phonotactic environment or orientation.

Articulatory constraints capture the tension between output and system orientation in primary and secondary phonotactic environments, respectively. Within the context of a relational understanding of segments and linguistic systems, the tension is provided as being greater or fewer systemic contrasts. In

---

[20] Traditional descriptions of intervocalic consonants as being in a weak position refer to noted propensity for these consonants to elide or undergo lenition. This does not correspond to my use of secondary as a positional description, as this term relies on the degree of articulatory ease or difficulty (assuming the gesture is, indeed, made) and perceptual salience.

essence, a segment's articulatory integrity specifies what it is and is not; within this definition, variation is allowed to the greatest extent possible that does not result in a violation of the systemically defined integrity. The following constraints capture the tension between Hyper and Hypo speech modes:

> SPEND: in primary environments, use more contrasts (among those systemically available) to establish an output oriented articulatory segment.
>
> SAVE: in secondary environments, use fewer contrasts (among those systemically available) to establish a system oriented articulatory segment.

SPEND and SAVE reflect the observations of Lindblom (1990 and 1983) and allow for the inclusion of H&H in an output oriented, OT analysis. Because Hyper and Hypo speech are considered a continuum, the two constraints are co-ranked.

H&H also applies to the perceptual drive, albeit indirectly, as it provides a goal-oriented grammar for speech modes. The relational approach to perception provides that issues of discriminability or contrast must be considered from the point of view of the system and the perceptual segments that are defined within and from this organic context. The following constraints define the tension established in output and system oriented speech gestures, as they affect the perceptual system of contrasts and similarities:

> CONTRAST: In primary environments, provide as many systemically defined contrasts as possible to promote an output oriented perceptual segment.
>
> NEUTRALIZE: In secondary environments, neutralize as many systemically defined contrasts as possible to provide a system oriented perceptual segment.

As with SPEND and SAVE, CONTRAST and NEUTRALIZE are co-ranked with regard to perceptual output.

5.2.1. FAITHFULNESS WITHIN A RELATIONAL PHONOLOGY

Classical uses of OT assume that faithfulness or output fidelity to input specification plays a role of equal importance to that of markedness. In such analyses, faithfulness constraints compete with markedness in the determination of an output form. Depending on ranking, faithfulness may override markedness in the determination of an optimal output candidate. As with many other adaptations in the present work, the notion of segmental faithfulness is considerably modified within a relational phonology.

Because the segment is defined within an organic system—one which is wholly unconcerned with other organic systems—and because features are emergent, rather than universal, faithfulness to such features is a system specific phenomenon. Furthermore, because a segment is defined with regard to all others and owes this definition to the critical factor of differentiation, a great deal of variation is captured within gestural and acoustic specifications, as well as within the relational definition of a given segment. For that reason, satisfaction or violation of FAITH constraints takes on an entirely different role, in opposition to that usually ascribed it in more conventional OT studies. The essential role of faithfulness within a relational model is the elimination of potentially optimal candidates that do not share input specifications, i.e. that are differently mapped.

As with all other families of constraints, my treatment of faithfulness distinguishes between those constraints applicable to each of the two drives. Within a relational understanding of segments, faithfulness is provided as follows:

> FAITH(a): the output articulatory integrity of a segment should correspond to the input articulatory integrity of a segment.
>
> FAITH(p): the output perceptual integrity of a segment should correspond to the input perceptual integrity of a segment.

Articulatory faithfulness weighs against the inclusion of supplemental gestural characteristics or the omission of any of these. Likewise, perceptual faithfulness states that no perceptual contrasts or distances in the input should be lost, nor should there be any in the output that were not originally present in the input. As in more traditional approaches to faithfulness, it is possible to distinguish between different levels of input-to-output faithfulness. For example, articulatory output may distinguish between FAITH (articulatory target) and FAITH (articulator) in a given language and rank each of these differently. This is the case in New England English, as discussed in 5.4, below.

The reader will note that a particular candidate may incur multiple violations of either faithfulness constraint. Each non-correspondent gestural or acoustic component of an output will result in a FAITH violation. All other things being equal, the optimal output will mirror the input to the greatest extent possible. Such one-to-one correspondence is not always possible, however. Analysis of Amsterdam Dutch variation presents one example where no possible candidate is selected by the articulatory and perceptual grammars without at least one relevant faithfulness violation; discussion of this is reserved for 5.3.

5.3. Allophony in Amsterdam Dutch: [ʀ] and [ɹ]

The Amsterdam Dutch /R/ is defined in Chapter Four in such as way as to encompass all variants of this sound. Without exception, speaker N3 produced [ʀ] in word-initial and intervocalic environments and [ɹ] in word-final environments. These observations concur with Collins and Mees (1996: 199-200), as well as with the dialectal and sociolinguistic studies of van de Velde (1994: 32-34), van Reenen (1994: 57-59), and de Schutter and Taeldeman (1994: 83-84). The latter works assert that the pre-vocalic uvular/velar versus post-vocalic bunched articulations are common to most dialects situated north of the rivers Waal, Maas, and Rhine.

Speaker N3 produced morphologically related tokens (e.g. *haren* – *haar*, *boeren* – *boer*, etc) of [ʀ] and [ɹ], strengthening traditional analyses of the two variants as being allophones. This alternation is expressed with traditional notation as below:

/R/ $\rightarrow$ [ʀ] / $_{\text{word}}[$___, V___V

/R/ $\rightarrow$ [ɹ] / ___$]_{\text{word}}$

Derivational phonologies are able to describe this case of variation, but can do little to explain it. Any derivational model must begin by first establishing which of the two allophones constitutes the underlying phoneme. An approach using feature geometry must also begin with a predetermined phoneme—the choice of which would be arbitrary—and capture variation on two planes: place of articulation, and manner of articulation. I am aware of no model of feature geometry that looks at specific perceptual characteristics in its analysis. My approach describes this case of allophony, explains it from the point of view of a relationally defined, organic linguistic system, and predicts other instances of variation. This is accomplished with functionally grounded constraints and without arbitrary underlying forms.

In order to describe allophonic variation using a relational theory, the segmental integrity of each of the two variant rhotics must first be understood. As with all other segments presented in this dissertation, opposition is maintained between gestural and acoustic specifications.

The articulatory grammars of each rhotic allophone are given in Table 5.1. From an articulatory point of view, these segments are quite different, but still fit within the larger picture of inter-linguistic rhotic integrity provided in 4.5.1. Word-initial and intervocalic [ʀ] is accomplished with the tongue body whose distal target is the uvula (the velum is a secondary or indirect target of this gesture). The tongue is non-contracted or flat for this allophone, but the gesture is periodic: the movement of articulators during pulmonic egression results in multiple open-close movements of the uvular target. Word-final [ɹ] is similar to word-initial [ʀ] by virtue of a shared tongue body articulator. The target of articulator movement is the palate, and the tongue is in a bunched contour. No periodicity is noted in word-final allophones.

	Articulator	Target	Art. Shape	Periodicity
ʀ	tongue body	uvula (velum)	flat	yes
ɹ	tongue body	palate	bunched	no

Table 5.1. Gestural specification for Amsterdam Dutch [ʀ] and [ɹ].

Acoustic specification in Amsterdam Dutch is provided according to the method outlined in 4.3 and provided in Table 5.2. Acoustic regularization for quality and quantity shows no qualitative contrast between the rhotic allophones, save for periodicity in the case of [ʀ]. Both [ʀ] and [ɹ] are characterized by the presence of formant peaks. Two qualitative distinctions are noted, however. The first of these is a relative raising of F1, from a range of 400-510 Hz in word-initial and intervocalic environments to 650-700 Hz in word-final environments. This variation is most likely attributed to word-final devoicing and is not significant: note that both English speakers, who regularly produced a similar rhotic in all target environments, produced tokens with F1 of approximately 350-500 Hz (Table 4.2). A more important perceptual distinction is the difference in F3: this is greater for [ʀ] than for [ɹ].

	Quality	Quantity (Hz)
ʀ	formants, periodicity	400-510, 1485-2085, 2700-2800
ɹ	formants	650-700, 1150-1800, 1735-2250

Table 5.2. Acoustic specification for Amsterdam Dutch [ʀ] and [ɹ].

These specifications provide the articulatory and perceptual integrity of each of the Amsterdam Dutch rhotic allophones. Articulatory and perceptual integrity for each rhotic allophone are given as follows:

> **[ʀ] (articulatory)**: "≠ /s, z/ (articulator, target, shape, periodicity), /ʃ/ (articulator, target, periodicity), /f/ (articulator, target, periodicity), /x/ (target, periodicity), /l/ (articulator, target, periodicity); = /ʃ/ (shape), /x/ (articulator, target, shape), /l/ (shape)."
>
> **[ʀ] (perceptual)**: "≠ /s, z/ (3 quantities, 3 qualities), /ʃ/ (2 quantities, 3 qualities), /f/ (3 quantities, 3 qualities), /x/ (2 quantities, 3 qualities), /l/ (1 quantity, 3 quality); /ʃ/ (1 quantity), /x/ (1 quantity), /l/ (2 quantities)."
>
> **[ɹ] (articulatory)**: "≠ /s, z/ (articulator, target, shape), /ʃ/ (articulator, shape), /f/ (articulator, target), /x/ (target, shape), /l/ (articulator, target, shape); = /ʃ/ (target), /x/ (articulator)."
>
> **[ɹ] (perceptual)**: "≠ /s, z/ (3 quantities, 3 qualities), /ʃ/ (3 quantities, 3 qualities), /f/ (3 qualities, 3 quantities), /x/ (1 quantity, 3 qualities), /l/ (1 quantity, 1 quality); = /x/ (2 quantities), /l/ (2 quantities, 2 qualities)."

The reader will note that the integrity of both rhotic allophones fits into the inter-linguistic relational integrity proposed in 4.3.

Returning to the constraint based descriptions of system interaction of 4.4, allophonic grammars of each allophone are generated, as provided in Tables 5.3 and 5.4, below.

	*HOLD (2)	*HOLD (1)	REUSE(cc)	REUSE(all)
[ɹ]		*	*	*
[ʀ]	*	*	*	*

Table 5.3. Articulatory grammar of Amsterdam Dutch rhotics [ʀ] and [ɹ].

Each rhotic allophone results in the same number of REUSE(cc) and REUSE(all) violations: [ɹ] due to a bunched articulator contour and [ʀ] due to a uvular articulatory target. The main distinction between these allophones is the number of *HOLD(2) violations: the periodic [ʀ] violates this constraint, whereas the [ɹ] satisfies it. All other things being equal, [ʀ] is a less optimal articulatory candidate, given its relation to the greater Amsterdam Dutch system, as this segment results in more systemic complexity.

The perceptual grammar of Amsterdam Dutch rhotics tells a different story about allophonic variation between [ʀ] and [ɹ], as seen in Table 5.4. Quantitative distinctions between the two allophones are not significant: the word-initial and intervocalic rhotic results in four *MINDIST(2) and two *MINDIST(3) violations, whereas the word-final allophone violates each of these constraints three times. Given this lack of significance, *MINDIST constraints are ignored in all further discussion of Amsterdam Dutch allophony. Qualitatively, [ʀ] proves more satisfactory and does not violate any MAINTAIN constraints, whereas [ɹ] violates MAINTAIN(3), due to similarities between this segment and /l/. The word-final rhotic is also qualitatively similar to vowels. Perceptually, [ʀ] is a better candidate than [ɹ]: it is more distinct within the system of Amsterdam Dutch similarities and oppositions.

	Ma(1)	Ma(2)	Ma(3)	*MD(3)	*MD(2)	*MD(1)
ɹ			*	***	***	*****
ʀ				**	****	*****

Table 5.4. Perceptual grammar of Amsterdam Dutch rhotics [ʀ] and [ɹ]. MAINTAIN and *MINDIST are abbreviated Ma and *MD, respectively.

The tension between the two drives relevant to speech production is exemplified by rhotic allophony in Amsterdam Dutch. The articulatory grammar favors [ɹ], which results in less system specific articulatory effort. By contrast, [ʀ] is favored by the perceptual grammar: the integrity of this segment provides that it is more distinct from all other Amsterdam Dutch continuants.

The articulatory segmental integrity of the rhotic allophones of Amsterdam Dutch show [ʀ] to be a higher cost member, when cost is defined as the increase of systemic gestural components. H&H derived constraints presented in 5.2, above, integrate system-external principles of articulation and perception into discussion of segmental behavior, in this case allophony. Table 5.5 provides the articulatory output of Amsterdam Dutch for two tokens.

Because /R/ allows for inclusion of all variants, as defined by its relation to the continuant system of Amsterdam Dutch, all allophones are both conclusive faithful (inclusive) and unfaithful (additive) to the input. This is shown with a violation of the relevant faithfulness constraint. The optimal candidate for both inputs satisfies the higher-ranking SAVE / SPEND constraints, even if greater effort is the result (in this case, violation of *HOLD). This is the case of [ʀux]:

because the rhotic is in a primary phonotactic environment, considerations of systemic efficiency are secondary to those of the output oriented speech mode.

/Rux/, /ha:R/	SAVE	SPEND	FAITH(a)
⇒ [ʀux]			**
[ɹux]		*!	***
[xux]			****!
[ha:ʀ]	*!		**
⇒ [ha:ɹ]			***
[ha:x]			****!

Table 5.5. Articulatory output of Amsterdam Dutch, /Rux/ and /ha:R/.

Note that FAITH(a) is a general faithfulness constraint and does not distinguish between articulator, target, articulator contour, or periodicity. Faithfulness plays a critical role, even if it is dominated by the co-ranked SAVE / SPEND, inasmuch as it prevents selection of [x] or any other systemically defined candidate. Potential candidates [xux] and [ha:x] satisfy markedness constraints, but present four violations of FAITH(a), the result of non-faithful target (lacking uvular and palatal components), tongue contour (lacking bunched contour), and periodicity components of [x].

Articulatory output for Amsterdam Dutch presents one output example where no candidate—optimal or suboptimal—fully satisfies faithfulness. This is the result of relational definition and specification of /R/, which is accomplished from the bias of systemic, discriminatory relations. Because the rhotic phoneme is defined in such a way as to encompass all variants, all allophonic manifestations of this segment are lacking in at least one of the gestural specifications. Articulatory output necessarily leads to FAITH(a) violation for all candidates.

Functionalism presumes any speech act to have as its goal the reception of an acoustic signal. Variation must therefore be considered from a perceptual point of view as well. Table 5.4 demonstrates that [ʀ] is a more distinct variant of /R/ than is [ɹ], as it provides maximal contrast with regard to all other segments of the Amsterdam Dutch system. In terms of perception, [ʀ] is easier to distinguish and [ɹ] is more confusable. Table 5.6 shows the result of CONTRAST and NEUTRALIZE constraints, ranked with regard to FAITH(p).

/Rux/, /ha:R/	CONTRAST	NEUTRALIZE	FAITH(p).
⇒ [ʀux]			*
[ɹux]	*!		
[xux]			****!
[ha:ʀ]		*!	*
⇒ [ha:ɹ]			
[xux]			*****!

Table 5.6. Perceptual output of Amsterdam Dutch /Rux/ and /ha:R/.

The optimal output for /Rux/ is [ʀux]: it results in no systemic MAINTAIN violations and no system-external perceptual violations. The sub-optimal candidate [ɹux] violates CONTRAST, as it does not provide maximal contrast vis-à-vis other system participants. The optimal candidate for the input /ha:R/ is [ha:ɹ]. Although it violates MAINTAIN(3), it satisfies all system-external constraints.

Two particularities with regard to faithfulness are noted in the perceptual output candidates and their satisfaction and/or violation of FAITH(p). The first of these involves candidates including [x]: these do not result in any markedness violations, as there is no relative promotion of or neutralization of contrasts. However, outputs containing the velar fricative violate FAITH(p) no fewer than four times, due to three non-congruent quantities and qualitative specification for dispersion, and represent sub-optimal candidates. The second observation to be made in regard to faithfulness is the opposition between perceptual and articulatory outputs. Whereas the latter optimal output necessarily engenders one (or more) faithfulness violations, the same is not true of perceptual candidates. The palatal approximant [ɹ] contains a perceptual correspondent to all acoustic, input specifications of /R/ and satisfies FAITH(p) at all levels, even if onset candidates violate higher ranking CONTRAST.

Amsterdam Dutch constitutes one example where markedness constraints are higher ranked with regard to faithfulness constraints. This system demonstrates that, given a relational definition of segments, the optimal output segment will result in the least amount of articulatory effort while providing a minimum of systemic discriminability. This approach also predicts that other instances of variation, such as the alternation between light and dark /l/ or vowel reduction, can be predicted using the same series of constraints.

5.4. Rhotic reduction: New England English [ɹ] and [ø]

The particularities of rhotics and rhotic variation serve as a primary phonological distinction for New England English from North American Standard English. These forms of English are spoken in the area bordered by the Connecticut River in the west, the Appalachian Range in the north, and the Saint Croix River in the east. This definition excludes the regional English of Northern Maine and New Hampshire, as well as those of Vermont, and the western counties of Massachusetts and Connecticut. In opposition to Standard American English, New England dialects regularly reduce rhotics in coda positions, where they surface as a schwa-like vowel. This so-called "rhotic drop" is a also common feature of nonstandard American English spoken in the coastal regions of the southeastern states and the New York City metropolitan area. Other dialects of English, namely those of Southern England, New Zealand, and Australia, also exhibit coda rhotic reduction or elision.

Previous analyses of rhoticization and rhotic variance in New England English have focused on rather generally delimited forms of the language and have typically sought to explain the assumed loss of /ɹ/ in word-final environments and in inflected suffixes. Many studies also address the so-called "hypercorrective r," a form of epenthesis where /ɹ/ surfaces in word-final environments lacking the rhotic in underlying or input forms.[21] McCarthy (1991)

[21] The hypercorrective /ɹ/ is attributed to linguistic insecurity in most sociolinguistic studies, few of which note that this phonological behavior is very similar to the "intrusive R" of British Received Pronunciation (RP). I do not treat this subject here in either sociolinguistic or

evaluates rhotic elision and insertion in Eastern Massachusetts English as a case of differential rule ordering. Elision is the result of phonotactic conditioning, whereas insertion is the product of a linking mechanism. McMahon (2000) treats the question of historical /r/ weakening as a lenition process, but does not specifically address synchronic alternation in the traditionally grouped rhotic-dropping dialects of English.

Analyses presented in this dissertation focus exclusively on the regional form of English spoken in southern coastal Maine (York and Cumberland counties), where the author has numerous familial and personal contacts. In addition to the above-mentioned rhotic reduction, these regional forms also show considerable lowering of non-high vowels and relatively exaggerated vocalic laxness in unstressed positions. The reader will note that, with the exception of rhotics and rhotic variants, the consonant system of New England English does not differ from that of Standard American. Likewise, no consonant—continuant or otherwise—shows variant behavior similar to that of the rhotic.[22]

Rhotic reduction in this form of New England English consists of /ɹ/ alternating with a vowel in all coda positions. Following non-high vowels, most notably /a/, this takes the form of a lengthened vowel. In other instances, i.e. following high, unrounded vowels, the rhotic surfaces as a neutral, schwa-like vowel preceded by the homorganic, postvocalic glide [j]. The following are

---

phonological regards, but do wish to note that rhotic behavior in New England English is essentially the same phenomenon as that of standard varieties of British English.

22 The examples provided in this section are presented from personal experiences and interviews conducted by the author and correspond to all available linguistic literature. Description of reduced forms as having a schwa-like vowel component correspond to the derivational models of Donnegan (1993).

examples of rhotic reduction, also giving counterexamples of the segment's maintenance in onset position:

car [ka:], *barn* [ba:n], *martin* [ma:tɪn], *Karen* [kæɹɪn]

here [hijɘ], *deer* [dijɘ], *earshot* [ijɘʃɑt], *hearing* [hiɹɪn].

In addition to these instances of rhotic elision, /ɹ/ is regularly reduced in polysyllabic words ending in *–er* (inflected and otherwise). In all such words, the rhyme of Standard American ([ɹ̩]) surfaces as a schwa-like vowel, as in the below examples:

spider [spidɘ], *faster* [fæstɘ], *bigger* [bɪgɘ].

A traditional analysis of rhotic-schwa/vowel alternation is described using slash-dash notation as follows:

/ɹ/ → [a] / V [+ low] ____

→ [ɘ] / V [- low] ____, ____] $_{\text{word}}$

→ [ɹ] / all others.

The rhotic alternates with either a schwa-like vowel or the homorganic [a] in coda positions. In all other environments, the rhotic surfaces as [ɹ].

A relational approach to the question of rhotic reduction looks to the variant forms—most notably the schwa-like vowel—and asks what loss of gestural and/or acoustic specification is implied by this variance. I use the term neutral segment (abbreviated as [ø]) with reference to the reduced form of /ɹ/ in both vocalic environments, avoiding potential confusion with other schwa-like segments that result from vowel reduction. I assume that this neutral segment is relationally defined in English as being a minimal gesture involving the tongue body (the vowel articulator) and having no further gestural specification. I further assume that its acoustic specification is defined as merely having formants, whose acoustic quantity is environmentally determined. This resolves the question of [ɘ] and [aː] alternation. If the neutral segment is defined as having no internal, relational integrity, it will inherit this from a surrounding segment or, when this is not possible (e.g. adjacent to an occlusive), it will default to neutral. I ignore instances of gliding and low vowel lengthening, assuming that these are two manifestations of the Obligatory Contour Principle (OCP), and ascribe both types of reduction to a singular phenomenon. Both gliding and low vowel lengthening are easily resolved with the markedness production and faithfulness perceptual constraints not included in the present analysis.[23]

As in 5.3, a primary task in explaining New England rhotic variation from a relational bias is the formalization of articulatory and perceptual specifications for [ɹ] and the neutral segment. Gestural specification is presented in Table 5.7, below. Note that considerations of periodicity are irrelevant: this is not a feature of the system in question and cannot be a part of any analysis of variance. As above, my analysis considers the neutral segment to be gesturally specified as having only the tongue body as its articulator, with no precise target or articulator shape being implied in the gesture. This reflects a schwa-like articulator specification.

[23] Boersma (1998: 415-440) formalizes OCP constraints within his functionalist grammar. Rather than reiterate this excellent analysis, I assume that gliding and schwa-[a] alternation are the products of his *GESTURE and PARSE constraints and refer the reader to this work for further discussion and implications.

	Articulator	Target	Art. Shape
ɹ	tongue body	palate	bunched
∅	tongue body	n/a	n/a

Table 5.7. Gestural specification for the New England English rhotics [ɹ] and [∅].

Distinction between [ɹ] and [∅] derives from the lack of tongue specification for articulatory target and contour, with regard to the latter variant.

The present analysis of New England rhotic variation provides that the acoustic quality of the neutral segment is defined by clear formants, but that quantity—or the exact frequency of these formants—is not part of its specification. These considerations are expressed in Table 5.8, below:

	Quality	Quantity (Hz)
ɹ	formants	350-650, 1000-1700, 1800-2200
∅	formants	n/a

Table 5.8. Acoustic specification of New England English rhotics [ɹ] and [∅].

The neutral segment's perceptual specification is such that the phonotactic environment of its instantiation will provide acoustic quantity—specific formant measurements—but that these quantities are not relationally significant.

These considerations are integrated into the articulatory and perceptual integrity of both rhotic allophones, as given below:

[ɹ] (articulatory): "≠ /s, z/ (articulator, target, shape), /ʃ/ (articulator, shape), /f, v/ (articulator, target, shape), /ð, θ/ (articulator, target, shape), /l/ (articulator, target, shape); = /ʃ/ (target)"

[ø] (articulatory): "≠ /s, z/, /ʃ/, /f, v/, /ð, θ/, /l/ (articulator, target, shape)"

[ɹ] (perceptual): "≠ /s, z/ (3 quantities, 3 qualities), /ʃ/ (3 quantities, 3 qualities), /f, v/ (3 quantities, 3 qualities), /ð, θ/ (3 quantities, 3 qualities), /l/ (1 quantity, 1 quality); = /l/ (2 quantities, 2 qualities)"

[ø] (perceptual) : "≠ /s, z/, /ʃ/, /f, v/, /ð, θ/, /l/ (3 qualities, 3 quantities); = /l/ (3 qualities)."

As with discussion of variation in Amsterdam Dutch, considerations of gestural and perceptual specification are formulated in constraint based grammars. These consider the interaction of *HOLD and REUSE constraints, showing how variants respond to the need for gestural economy specified in Chapter 2. Table 5.9, below, presents the articulatory grammar of each New England variant.

	*HOLD (2)	*HOLD (1)	REUSE(cc)	REUSE(all)
ɹ		*	**	*
ø		*	*	

Table 5.9. Articulatory grammar of New England rhotic allophones [ɹ], [j], and [ø].

Two observations are made of the articulatory grammar of rhotics in this system. Firstly, the neutral segment has only limited gestural specification—since no target or tongue shape are implied by it—and results in only one systemically defined violation of REUSE(cc) (tongue body articulator). By contrast, [ɹ] violates REUSE(cc) twice (tongue body articulator and bunched articulator shape), and REUSE(all) once, (bunched shape). Within the context of a relationally defined system, [ø] is a more efficient articulatory variant.

The perceptual grammar of New England rhotics reveals information about the relational discriminability of these variant segments. Here, *MINDIST is of lesser importance than is MAINTAIN, with regard to the relational contrasts implied by the different segments. These are presented in Table 5.10, below:

	Ma(1)	Ma(2)	Ma(3)	*MD(3)	*MD(2)	*MD(1)
ɹ		*	*	****	****	*****
ø	*	*	*			

Table 5.10. Perceptual grammar of New England rhotic allophones.

Considered from a perceptual point of view, [ɹ] is the best candidate, as it presents the greatest number of systemically defined contrasts. The neutral segment, while it does not present any violations of *MINDIST, fails to satisfy MAINTAIN by default, as the positive provision of this constraint provides for its violation by virtue of a lack of qualitative contrast with regard to other, qualitatively similar segments. Within the continuant consonant inventory, there is potential for perceptual confusion with /l/. More importantly, discriminability with regard to the larger system of vowels is extremely reduced, making the neutral segment a sub-optimal perceptual candidate.

H&H derived markedness constraints distinguish between primary (onset) and secondary (coda) phonotactic environments and state that the articulatory output will provide for the greatest amount of systemically defined articulatory ease. For this analysis, it is necessary to distinguish between different faithfulness constraints: FAITH (articulator) states that there must be an input-output correspondence for the gestural articulator parameter; FAITH (target and shape) provides output correspondence for target and articulatory contour specifications. FAITH (articulator) is ranked higher than either markedness constraint. This is necessary to prevent full rhotic elision, as shown in Table 5.11, below.

/kaɹ/, /kɛɹɪn/	FAITH (articulator)	SAVE	SPEND	FAITH (target, shape)
[kaɹ]		*!		
⇒ [ka:]				**
[kal]	*!	*		**
[ka]	*!			**
⇒ [kæɹɪn]				
[kæ:ɪn]			*!	**
[kælɪn]	*!			**
[kæɪn]	*!		*	**

Table 5.11. Articulatory output of New England English tokens /kaɹ/ and /kɛɹɪn/.

Articulatory output provides that, in secondary environments, the rhotic will be reduced, but that this segment will be protected in primary environments. Candidates containing /l/ (or any other systemically related segment) are deemed sub-optimal, due to violations of faithfulness constraints. Candidates where the entire rhotic gesture is lost are also sub-optimal, regardless of any markedness satisfaction, due to violations of Faith (articulator). Here, faithfulness plays an important role and prevents the loss of the metrical slot occupied by the rhotic in lexical input.

The perceptual grammar also demonstrates the importance of faithfulness constraints, although it is not necessary to distinguish between qualitative and quantitative perceptual faithfulness in the New England English output grammar. Table 5.16 presents two cases of perceptual output, using the same examples as in 5.15, above.

/kaɹ/, /kɛɹɪn/	CONTRAST	NEUTRALIZE	FAITH (p)
[kaɹ]		*!	
⇒ [ka:]			***
[kal]		*!	***
[ka]			****!
⇒ [kæɹɪn]			
[kæ:ɪn]	*!		***
[kælɪn]			***!
[kæɪn]	*!		****

Table 5.12. Perceptual output of New England English tokens /kaɹ/ and /kɛɹɪn/.

Here, the optimal candidate satisfies FAITH(p) at all levels, while also avoiding violations of NEUTRALIZE or CONTRAST. Despite FAITH(p) violations, coda reduction provides more neutralization, whereas onset maintenance allows for greater perceptual discriminability.

5.5. Free variation in French: [ʁ̝] and [ʁ̞]

Surprisingly little mention of French rhotics has been made in either phonological or phonetic literature. One exception to this is seen in Tranel (1987), who describes the /ʁ/ as having uvulo-velar or velar place of articulation and either fricative or approximant manner of articulation. In his analysis, onset and intervocalic rhotics tend to be more fricative than coda—especially pre-occlusive—tokens. Alternation between fricatives and approximants is assumed to be a case of free variation, i.e. responsibility for the choice is assigned to individual speakers. Tranel further states that, regardless of which manner of articulation is employed, rhotics are devoiced in word-final instantiations and in adjacency to voiceless consonants. As with all previous discussion, I ignore voicing variation in this analysis.

Description of the French rhotic /ʁ/ has thus far been as inclusive as possible, allowing for the greatest degree of gestural and perceptual variation. Little predictable variation, save for the usual interference of vowel-to-consonant coarticulation, is noted in speaker data of either subject F5 or F6, as presented in Chapter Three. In word-initial environments, the rhotic always displays more formant dispersion. Word-final rhotics show, by contrast, less dispersion, although there is no means to provide for the absolute qualification of variation. Intervocalic tokens are the most variant forms of all. For both speakers, there is clear impressionistic and measured evidence of frication in tokens; however, this varies from token to token and between the two speakers. The lack of clear patterns and of absolute distinctions between more fricative rhotics ([ʁ̝]) and less fricative rhotics ([ʁ̞]) would seem to support the free variation hypothesis and presents a challenge for relational phonology.

Relational definition derives first from the gestural and acoustic specifications of each rhotic variant. Tables 5.13 and 5.14 demonstrate that these specifications do not allow for distinction between the two rhotics based on relational considerations:

	Articulator	Target	Art. Shape
ʁ̞	tongue body	velum/uvula	flat
ʁ̝	tongue body	velum/uvula	flat

Table 5.13. Gestural specification for the French rhotics [ʁ̞] and [ʁ̝]

Given the minimally distinctive gestures implied by the production of [ʁ̞] and [ʁ̝], gestural specification does not provide for a motivated, relational distinction.

	Quality	Quantity (Hz)
ʁ̞	dispersion, formants	460-965, 1220-1840, 2800-3425
ʁ̝	dispersion, formants	460-965, 1220-1840, 2800-3425

Table 5.14. Acoustic specification for the French rhotics [ʁ̞] and [ʁ̝].

Acoustic specification also denies any relational distinction between the two rhotics. This does not deny the validity of observations of the differences, but simply states that these differences are not important to the segmental integrity of the two rhotics, as provided below:

> **[ʁ̝] and [ʁ̞] (articulatory)**: "≠ /z/ (articulator, target, shape), /ʒ/ (articulator, target), /v/ (articulator, target), /l/ (articulator, target); = /ʒ/ (shape), /l/ (shape)."

> **[ʁ̝] and [ʁ̞] (perceptual)**: "≠ /z/ (3 quantities, 3 qualities), /ʒ/ (2 quantities, 3 qualities), /v/ (3 quantities, 3 qualities), /l/ (1 quantity, 3 qualities); = /ʒ/ (1 quantity), /l/ (2 quantities)."

Variation in this instance is not relationally important and, thus, cannot be treated by relationally biased articulatory and perceptual grammars, such as those given in Tables 5.15 and 5.16, below.

	*HOLD (2)	*HOLD (1)	REUSE(cc)	REUSE(all)
ʁ̝		*	**	
ʁ̞		*	**	

Table 5.15. Articulatory grammar of French rhotics [ʁ̝] and [ʁ̞].

	Ma(1)	Ma(2)	Ma(3)	*MD(3)	*MD(2)	*MD(1)
ʁ̞				**	***	****
ʁ				**	***	****

Table 5.16. Perceptual grammar of French rhotics [ʁ̝] and [ʁ̞].

It is quite natural, therefore, that articulatory and markedness constraints, as well as the corresponding faithfulness constraints as they are provided in section 5.2 have no effect on the output of relationally defined inputs. Tables 5.17 and 5.18 demonstrate that SAVE and SPEND, as well as CONTRAST and NEUTRALIZE, are unable to promote or forbid one or the other assumed variants.

Like analyses of variation in Dutch and English, faithfulness constraints prevent the selection of candidates whose articulatory and perceptual specifications do not correspond the respective input parameters. In the tableaux below, I have included only one example of an unfaithful candidate that is eliminated by FAITH; any other systemically related segment could be included in its place.

/ʁa/, /paʁ/	SAVE	SPEND	FAITH(A)
⇒ [ʁ̝a]			
⇒ [ʁ̞a]			
[la]			***!
⇒ [paʁ̝]			
⇒ [paʁ̞]			
[pal]			***!

Table 5.17. Articulatory output for French tokens /ʁa/ and /paʁ/.

/ʁa/, /paʁ/	CONTRAST	NEUTRALIZE	FAITH(p)
⇒ [ʁ̝a]			
⇒ [ʁ̞a]			
[la]			****!
⇒ [paʁ̝]			
⇒ [paʁ̞]			
[pal]			****!

Table 5.18. Perceptual output for French tokens /ʁa/ and /paʁ/.

Articulatory and perceptual output fails to distinguish between two optimal candidates and provides that either may be selected in primary or secondary environments. The lack of distinction between more and less fricative rhotics in the output candidate set would seem to result in a failure of relational phonology and present an unanswerable challenge to relational approaches to segment definition. If a relationally approach to rhotics and rhotic classhood cannot explain this variation, has relational phonology failed to fulfill its mission, i.e. to account for linguistic form and behavior?

I wish to firmly state that it has not. Moreover, I assert that such an apparent failure on the part of relational phonology only serves to strengthen this model, by making it clear what phonology should and can do, and what it cannot. Rather than dwell on relational phonology's failure to provide definitive answers to this question, I see the lack of an answer to the question of French rhotic variation as one of the principal strengths of this theory. It is useful to recall the definition of segments within relational phonology: these are defined with only those features that are present in the organic system and only as a function of significant difference and similarity of one system member with regard to all others. Production of /ʁ/ in French as a fricative or as an approximant has no effect on any of the systemic differences or similarities that are used to provide this segment its integrity, i.e. its place within the system of French sound segments. No variant of [ʁ], whether fricative or approximant, will be confused with /l/, /z/, or any other continuant consonant of French. This instance of variation cannot be explained by relational phonology or by any other phonological model. Phonology may describe such variation and explain how it is allowed to take place within the system, but it cannot predict one output or another, because this type of variation does not affect the phonological dynamic.

Assuming that no variation is truly free, i.e. without some form of catalyst, how can the differences between rhotics be explained? I wish to assert that the explanation for variance such as that seen in French is a question for phoneticians and for linguists focusing on the communicative context. Lindblom (1990) makes room for social and communicative constraints within his dynamic of intraspeaker interaction (418-419). These constraints concern the context of speech and the impetus for the communicative act, although no formalizations as to the effect of different contexts on either output or system orientations are offered. Clearly, there is need for further investigation into the effects of such social and communicative constraints on output.

5.6. Conclusion

In describing and explaining instances of rhotic variation, the strengths and weaknesses of relational phonology and of the construction of a rhotic relational class are made clearer. The task of relational phonology consists in providing a model for the segmental integrity—both articulatory and perceptual—of different segments and, in cross-linguistic studies, of the congruence of such segmental integrity. Relational phonology is able to explain instances of allophony where the output form of an input phoneme results in one or more violations of faithfulness constraints, i.e. a change in the gestural and/or acoustic specification of the original segment and its articulatory and perceptual grammar. Two such types of allophony are presented above. Relational phonology is not able to explain the output form of segments where those outputs do not result in a modification of an input specification. This is seen in the assumed free variation of the French /ʁ/. Relational phonology and the notion of segmental integrity may, however, state why such variation is allowed to occur.

Treatment of rhotic variation within a relationally defined phonology raises questions about the role of phonology and underscores the need for enlargement of phonological research and models. Future work in relational phonology must strive to include in the scope of its evaluation and explanation such diverse linguistic factors as the social context of communication and the biology of communicative actors, both of which have been heretofore deemed irrelevant to phonology as a discipline. It is hoped that the analyses here, however tentative some of the conclusions may be, will serve as an introduction to further study in this regard.

CONCLUSION

This dissertation began by questioning the validity of the term rhotic and a handful of studies that have attempted to motivate rhotic classhood. In the preceding chapters, I have demonstrated that the disparate articulatory and perceptual specification of individual rhotics denies any motivation of a structural classhood and have offered an innovative theory for the relational integrity or definition of segments. This dissertation describes rhotic as being a relational class of sounds and provides a functionalist, relational understanding of phonological features and segments. I have also discussed linguistic variation, examining how a relational theory and its particular definition of phonological segments can account for particular rhotic behavior. These innovations derive from a bias towards the organic linguistic system, assuming that any discussion of sound segments must be grounded in observable linguistic reality and that such reality is subject to higher, supralinguistic requirements of the human communicative and biomechanical systems.

Relational phonology and the specific notion of relational integrity of segments present several challenges to traditional phonology. Firstly, it denies many of the fundamental bases of this discipline as being irrelevant—by virtue of their synthetic nature—and challenge the phonologist to begin her or his investigation from a more phonetic stance. Indeed, the first task for any question of relational phonology is the description of the organic linguistic system; this is an inherently phonetic enterprise. Evaluation of this description is, however, purely phonological, as it seeks to provide for discriminatory regularities of observed linguistic output. These regularities must be motivated and emergent, as must be any constraints posited of the system; once again, relational phonology demands a great deal of phonetic understanding on the part of its adherents.

A second, and perhaps more important distinction presented by relational phonology is its redefinition of the phonological segment. While the phonetic segment may indeed have implicit, tacit reality, the regularization of phonetic observations and the inclusion of the phonological segment—with its respective gestural and acoustic specifications—deny any a priori substance, such as those presented in many of the structuralist analyses of Chapters One and Two. The relational, phonological segment has no implicit structure, but gains structure or specificity by virtue of its dynamic, multiplex articulatory and perceptual distinctions with regard to a greater organic whole. In essence, the phonological segment, defined within a relational phonology, is nothing apart from its organic context. This is, indeed, a tremendous challenge for phonology. While providing for the relational classhood of rhotics, I have in essence denied any universal notion of segmental classhood.

The working theory of relational phonology, as outlined in this dissertation, has necessarily been limited in not only its focus but also its scope. Chapters Two and Four specifically mentioned these limitations with regard to the elaboration of perceptual specifications and to perceptual relations. Future work in relational phonology must incorporate a greater understanding of auditory and cognitive processes involved in perception and the influence of these on the perceptual drive. This will deepen our perspective on perception, viewed as a cognitive and auditory process and subject to numerous other constraints, and will provide a better understanding of how systems organize themselves with regard to the principles laid out in Chapter Two.

Future work on the question of rhotics will also look to other languages having other rhotics, as well as to instances of diachronic change. One particularly interesting example of the former is Portuguese, where phonetic and phonological literature attest to at least two different r-like sounds in contrastive distribution. How would relational phonology account for the special relation of these two sounds to each other, as well as their relation to the greater inventory of the language? Will relational phonology have to provide for a unified description of rhotics, i.e. distinction of these segments from all others on a primary level, as well as secondary distinction within the greater system? An example of historical change in rhotic specification is seen in the standard form of French. Here, apical rhotics were supplanted by the current dorsal rhotics during the seventeenth and eighteenth centuries. During a time of transition among middle and upper class speakers, there are many accounts of /r/ being confused with /z/ and /s/, as well as some historical evidence for the two segment's confusion (e.g. *chair* and *chaise*). How might a relational phonology account for diachronic change and for transitory phenomena such as this?

Discussion of future work using relational phonology (and hopefully contributing to its evolution and development) raises the question of what can and cannot be treated from this perspective. For the moment, I wish to respond solely with regard to phonology and questions of phonological interest, although I suspect relational principles might also be applicable to other domains of linguistic inquiry. Clearly, many types of analyses may be pursued from a relational perspective (e.g. segment typology and classification, sound change, allophony). Such studies benefit from a grounded, phonetically motivated conception of the phonological segment and from the relational definition of systems and system dynamics. Other phonological problems may be approached from an intrinsically relational manner, such as assimilation processes and elision or suppletion. Here, relational phonology can serve to better understand those elements of a phonological grammar that are undergoing particular processes; by defining the segments and/or systems in questions from an organic, relational perspective, a phonologist might best describe not only what is taking place, but how and why this phenomenon occurs (or is allowed to occur). As such, relational phonology would not replace, but enhance analysis using other models (notably, generative OT).

Final mention must be made of the question of falsification, both with regard to the specific analyses proposed in this dissertation and to the working theory of relational phonology, as well as the model proposed by it. While the nature of this theory does not allow for falsification in the traditional sense of the term, many shortcomings of this approach deserve mention in closing. Some of these are mentioned above and acknowledged throughout the dissertation, namely the limited scope of both the articulatory and perceptual drives. Others are obvious to phonologists working within more traditional, structurally biased theories. In each of these instances, objections may be raised. The psycholinguist will be disappointed by the lack of mention and allowance for such issues as trading relations or the normalization of auditory signals. For a phonologist using classic OT models, the denial of features and the highly questionable nature of classhood, viewed from the relational standpoint as provided in this dissertation, may make a relational phonology of little use. In both cases, criticism is both fair and welcome by the author. In evolving my own understanding of phonology and of the segments in focus in this work, I have not sought to provide ultimate, Aristotelian Truth, but a glimpse at one conception of a truth, as applied to the question at hand and as might be useful—with more work and contributions from colleagues—to the description, evaluation, and explanation of other linguistic phenomena.

It is clear that many questions remain unanswered with regard to both the development of relational phonology, in general, and to the notion of rhotic relational classhood, in particular. The theory and model presented in this dissertation leave open the door for analysis of any number of language forms and variations; this is perhaps the best indication of theoretic potential and of explanatory success.

APPENDIX A: EXPERIMENTAL STIMULI

English Stimuli

[ɹ]: rare, read, rude, daring, earring, touring, car, ear, pour

[s, z]: sat, seat, suit, pausing, easing, oozing, pause, ease, ooze

[l]: late, lead, lure, mauling, ceiling, cooling, awl, eel, spool

[ʃ]: shall, sheet, sure, caution, Venetian, cushion, gosh, leash, tush

[v, f]: vault, veal, voom, coughing, heaving, moving, suave, heave, move

[ð, θ]: that, these, thus, slather, seething, soothing, moth, seethe, booth

Dutch Stimuli

/R/ raar (‘strange’), ried (‘guess-PAST’), rug (‘back’), hare (‘hers’), dieren (‘animal-PLU’), boeren (‘farmer-PLU’), haar (‘her-POSS’), dier (‘animal’), boer (‘farmer’)

[l] laat (‘late’), lied (‘song’), lucht (‘air’), halen (‘get-INF’), zielen (‘soul-PLU’), boelen (‘quantity-PLU’), haal (‘get’), ziel (‘soul’), boel (‘quantity’)

[z, s] zaad (‘seed’), ziek (‘sick’), zucht (‘seek-PAST’), azen (‘ace-PLU’), kiezen (‘choose-INF’), bozen (‘angry-PLU’), daas (‘bent’), kies (‘choose-SING’), boos (‘angry’)

[v, f] vaag (‘vague’), vies (‘dirty’), voet (‘foot’), haven (‘harbor-PLU’), dieven (‘thief-PLU’), hoeven (‘able-INF’), kaf (‘ball’), dief (‘thief’), hof (‘court’)

[x, ɣ] gaaf ('give-PAST'), gierig ('miserly'), goed ('good'), agaat ('agate'), diegene ('that one'), hoge ('high-PLU'), haag ('hedge'), zeeg ('say-SING'), hoog ('high')

[ʃ] sjaat ('cat-DIMIN'), sjiek ('chic'), sjoel ('school'), kaasje ('cheese-DIMIN'), meisje ('maiden, miss'), poesje ('cat-DIMIN')

French Stimuli

[ʁ]: rat ('rat'), rit ('laugh-3S'), roux ('ruddy'), Arras, irriter ('irritate-INF'), touron ('nougat'), par ('by'), pire ('worst'), pour ('for')

[l]: la ('the-FEM'), lit ('bed'), loue ('rent-3S'), alla ('go-PAST-3S'), illimité ('unlimited'), Toulon, balle ('ball'), mille ('thousand'), coule ('flow-3S')

[z]: Zaïre, zizi ('pipi'), zoulou ('zulu'), bazar ('bazaar'), physique ('physique'), cousin ('cousin'), base ('base'), bise ('kiss'), douze ('twelve')

[ʒ]: j'ai ('I have'), j'irai ('I will go'), joue ('play-3S'), gagea ('conjecture-PAST-3S'), Fiji, toujours ('always'), cage ('cage'), tige ('stem'), bouge ('move-3S')

[v]: va ('go-3S'), vie ('life'), vous ('you-PL'), avare ('miserly'), civile ('civil'), pouvons ('able-1P'), bave ('salivate-3S'), cive ('shallot'), louve ('female wolf')

APPENDIX B: INDEX OF RECORDED STIMULI

Click highlighted words to access sound recording, waveform, spectrogram, and spectra for each of the stimuli tokens. This launches a java executable in either Netscape or Explorer (depending upon default browser settings).

A description of the experiment, subjects, stimuli, and recording methods is provided in Chapter Three. Recordings included below have been edited to reduce noise and eliminate background hum,[24] as well as boost higher frequencies.[25]

Subject E1 (English)
rare, read, rude, daring, earring, touring, car, ear, pour
sat, seat, suit, pausing, easing, oozing, pause, ease, ooze
late, lead, lure, mauling, ceiling, cooling, awl, eel, spool
shall, sheet, sure, caution, Venetian, cushion, gosh, leash, tush
vault, veal, voom, coughing, heaving, moving, suave, heave, move
that, these, thus, slather, seething, soothing, moth, seethe, booth

Subject E2 (English)
rare, read, rude, daring, earring, touring, car, ear, pour
sat, seat, suit, pausing, easing, oozing, pause, ease, ooze
late, lead, lure, mauling, ceiling, cooling, awl, eel, spool
shall, sheet, sure, caution, Venetian, cushion, gosh, leash, tush
vault, veal, voom, coughing, heaving, moving, suave, heave, move
that, these, thus, slather, seething, soothing, moth, seethe, booth

---

[24] Background noise was eliminated by first constructing a noise profile of non-speech portions of the recording. The peak value of token segments were examined, the smallest of which was determined to be the baseline; the FFTs of all token segments whose peak values were less than 33% above the baseline were then averaged to create the noise profile. The noise profile was subtracted from the original (unedited) FFT of each segment and the waveform was reconstructed using reverse FFT.

[25] Emphasis used the formula, $x(n) = x(n) - ax\,(n - 1)$, where $a = 0.95$.

Subject N3 (Dutch)
raar, ried, rug, hare, dieren, boeren, haar, dier, boer
laat, lied, lucht, halen, zielen, boelen, haal, ziel, boel
zaad, ziek, zucht, azen, kiezen, bozen, daas, kies, boos
vaag, vies, voet, have, dieven, hoven, kaf, dief, hof
gaaf, gierig, goed, agaat, diegene, hoge, haag, zeeg, hoog
sjaat, sjiek, sjoel, kaasje, meisje, poesje

Subject N4 (Dutch)
raar, ried, rug, hare, dieren, boeren, haar, dier, boer
laat, lied, lucht, halen, zielen, boelen, haal, ziel, boel
zaad, ziek, zucht, azen, kiezen, bozen, daas, kies, boos
vaag, vies, voet, have, dieven, hoven, kaf, dief, hof
gaaf, gierig, goed, agaat, diegene, hoge, haag, zeeg, hoog
sjaat, sjiek, sjoel, kaasje, meisje, poesje

Subject F5 (French)
rat, rit, roux, Arras, irriter, touron, par, pire, pour
la, lit, loue, alla, illimité, Toulon, balle, mille, coule
Zaïre, zizi, zoulou, bazar, physique, cousin, base, bise, douze
j'ai, j'irai, joue, gagea, Fiji, toujours, cage, tige, bouge
va, vie, vous, avare, civile, pouvons, bave, cive, louve

Subject F6 (French)
rat, rit, roux, Arras, irriter, touron, par, pire, pour
la, lit, loue, alla, illimité, Toulon, balle, mille, coule
Zaïre, zizi, zoulou, bazar, physique, cousin, base, bise, douze
j'ai, j'irai, joue, gagea, Fiji, toujours, cage, tige, bouge
va, vie, vous, avare, civile, pouvons, bave, cive, louve

REFERENCES

Azevedo, Milton M. 1981. *A Contrastive Phonology of Portuguese and English*. Washington (DC): Georgetown UP.

Banczerowski, Jerzy. 1987. Towards a Dynamic Approach to Phonological Space. *Studia Phonemica Posnaniensia* 1: 5-30.

Barry, William J. 1995a. Schwa vs. Schwa + /r/ in German. *Phonetica* 52: 228-235.

Barry, William J. 1995b. Variation in Schwa + /r/ in German. *Proceedings: International Conference of Phonetic Sciences, Stockholm* 2: 214-217.

Beckman, Jill N. 1998. *Positional Faithfulness*. Ph.D. Dissertation, University of Massachusetts at Amherst.

Bladon, R.A.W. and F. J. Nolan. 1977. A video-fluorographic investigation of tip and blade alveolars in English. *Journal of Phonetics* 5: 185-193.

Boersma, Paul. 2000. The odds of eternal optimization in OT. Rutgers Optimality Archive (http://roa.rutgers.edu/), ROA 429-1200.

Boersma, Paul. 1998. *Functional Phonology: Formalizing the interactions between articulatory and perceptual drives*. The Hague: Holland Academic Graphics.

Bonet, Eulàlia and Joan Mascaró. 1997. On the representation of contrasting rhotics. *Issues in the Phonology and Morphology of the Major Iberian Languages* (Martínez-Gil, Fernando and Afonso Morlaes-Fient, editors). Washington (DC): Georgetown UP.

Browman, Catherine P. and Louis Goldstein. 1990. Gestural specification using dynamically-defined articulatory structures. *Journal of Phonetics* 18: 299-320.

Browman, Catherine P. and Louis Goldstein. 1989. Articulatory gestures as phonological units. *Haskins Laboratories Status Report on Speech Research* SR-99/100: 69-101.

Brown, Adam. 1988. Linking, intrusive, and rhotic /r/ in pronunciation models. *Journal of the International Phonetic Association* 18, 2: 144-151.

Burzio, Luigi. 2000. Segmental Contrast Meets Output-to-Output Faithfulness. *The Linguistic Review* 17, 2-4: 367-384.

Burzio, Luigi. 1998. Multiple Correspondence. *Lingua* 104, 1-2: 79-109.

Carré, René and Pierre L. Divenyi. 2000. Modeling and Perception of 'Gesture Reduction.' *Phonetica* 57: 152-169.

Chomsky, Noam and Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.

Clemens, G. Nick. 1985. The geometry of phonological features. *Phonology Yearbook* 2: 225-252.

Clemens, G. Nick and Elizabeth V. Hume. 1995. The internal organization of speech sounds. *The Handbook of Phonological Theory* (John A. Goldsmith, editor): 245-306. Oxford: Blackwell.

Collins, Beverley and Inger M. Mees. 1996. *The Phonetics of English and Dutch*, 3rd Edition. Leiden: Brill.

Dart, Sarah H. 1991. Articulatory and Acoustic Properties of Apical and Laminal Articulations. *UCLA Working Papers in Phonetics* 79.

Delattre, Pierre. 1971. Pharyngeal features in the consonants of Arabic, German, Spanish, French and American English. *Phonetica* 23: 129-155.

Delattre, Pierre. 1965. *Comparing the Phonetic Features of English, French, German and Spanish: An Interim Report*. Heidelberg: Julius Groos.

Delattre, Pierre and Donald C. Freeman. 1968. A Dialect Study of American R's by X-ray Motion Picture. *Linguistics* 44: 29-68.

Diehl, Randly L. 2000. Searching for an Auditory Description of Vowel Categories. *Phonetica* 57: 267-74.

Diehl, Randy L. and Bjorn Lindblom. 2000. Explaining the Structure of Feature and Phoneme Inventories. *Speech Processing in the Auditory System* (S. Greenber, W.A. Ainsworth, A. Popper, and R. Fay, editors). New York: Springer Verlag.

Dixon, R. M. W. 1972. *The Dyirbal Language of North Queensland*. Cambridge (UK): Cambridge UP.

Donegan, Patricia. 1993. On the phonetic basis of phonological change. In *Historical Linguistics: Problems and Perspectives* (Charles Jones, ed.). London: Longman. 98-130.

Elert, Claes-Christian. 1970. *Ljud och ord I svenskan*. Stockholm: Almquist and Wiksell.

Flemming, Edward. 2001. Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology* 18: 7-44.

Flemming, Edward. 1995. *Auditory Representations in Phonology*, Ph.D. Dissertation. University of California at Los Angeles.

Gimson, A.C. 1970. *An Introduction to the Pronunciation of English*. New York: St. Martin's Press.

Holt, Lori L. and Keith R. Kluender. 2000. General Auditory Processes Contribute to Perceptual Accommodation of Coarticulation. *Phonetica* 57: 170-180.

Hume, Elizabeth and Keith Johnson. 2000. A Model of the Interplay of Speech Perception and Phonology. To appear in *The Role of Speech Perception in Phonology* (Elizabeth Hume and Keith Johnson, editors). Academic Press.

Jakobson, Roman. 1916/1963. *Essais de linguistique générale: les fondations du langage*. Paris: Minuit.

Jakobson, Roman. 1939/1962. Zur Struktur des Phonems. *Selected Writings Volume 1, Phonological Studies*. The Hague: Mouton.

Jakobson, Roman, C. Gunnar Fant and Morris Halle. 1967. *Preliminaries to Speech Analysis: The Distinctive Features and their Correlates*. Cambridge (MA): MIT Press.

Jespersen, Otto. 1890/1920. *Lehrbuch der Phonetik*. Leipzig: Teubner.

Jessen, Michael. 1994. Functional differences between consonants and vowels in German. *Linguistische Arbeiten: Universale phonologische Strukturen und Prozesse*. Tübingen: Niemeyer.

Jones, Daniel. 1922. *An Outline of English Phonetics*. New York: G. E. Stechert & Co.

Kaiser, L. 1958. *Het spreken in beeld: anatomisch en phonetisch kaartwerk.* Amsterdam: Noord-Hollandsche Uitgevers.

Kager, René. 1999. *Optimality Theory*. Cambridge (UK): Cambridge UP.

Keating, Patricia A. 1983. Comments on the jaw and syllable structure. *Journal of Phonetics* 11: 401-406.

Kent, Ray D. and Charles Read. 1992. *The Acoustic Analysis of Speech*. San Diego: Singular.

Kingston, John N. 1992. Phonetics and Phonology of Perceptually Motivated Articulatory Covariation. *Language and Speech* 35, 1-2: 99-113.

Kingston, John N. 1990. Articulatory binding. In *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (John Kingston and Mary Beckman, eds.). Cambridge: CUP. 406-434.

Kingston, John N. and Randy L. Diehl. 1995. Intermediate properties in the perception of distinctive feature values. *Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV* (B. Connell and A Arvantti, editors). Cambridge (UK): CUP.

Kingston, John N. and Randy L. Diehl. 1994. Phonetic Knowledge. *Language* 70, 3: 419-454.

Kohler, K.J. 1990. Segmental Reduction in Connected Speech in German: Phonological Facts and Phonetic Explanations. *Speech Production and Speech Modelling* (W. J. Hardcastle and A. Marchal, editors): 69-92. Dordrecht: Kluwer

Ladefoged, Peter. 1971. *Preliminaries in Linguistic Phonetics*. Chicago: U of Chicago Press.

Ladefoged, Peter. 1957. Use of Palatography. *Journal of Speech and Hearing Disorders* 22, 5: 764-774.

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.

Ladefoged, Peter and Zongji Wu. 1984. Places of articulation: an investigation of Pekingese fricatives and affricates. *Journal of Phonetics* 12: 267-278.

Ladefoged, Peter, Anne Cochran and Sandra Disner. 1977. Laterals and trills. *Journal of the International Phonetic Association* 7, 2: 46-54.

LaRiviere, C., H. Winitz and E. Herman. 1975. The distribution of perceptual cues in English prevocalic fricatives. *Journal of Speech and Hearing Research* 18: 613-622.

Liljencrants, Johan and Bjorn Lindblom. 1972. Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrasts. *Language* 48: 839-62.

Lindau, Mona. 1985. The Story of /r/. *Phonetic Linguistics: Essays in honor of Peter Ladefoged* (Victoria A. Fromkin, editor): 157-168.

Lindblom, Bjorn 1990. Explaining Phonetic Variation: a Sketch of the H&H Theory. *Speech Production and Speech Modelling* (W. J. Hardcastle and A. Marchal, editors). Dordrecht: Kluwer. 403-439.

Lindblom, Bjorn. 1986. Phonetic Universals in Vowel Systems. *Experimental Phonology* (J.J. Ohala and J. J. Jaeger, editors).

Lindblom, Bjorn. 1983. Economy of Speech Gestures. *The Production of Speech* (Peter F. MacNeilage, editor). Heidelberg: Springer. 217-245.

Lindblom, Bjorn et al. 1995. Is sound change adaptive? *Rivista di Linguistica* 7, 1: 5-37.

Lipski, John M. 1990. Spanish Taps and Trills: Phonological Structure of an Isolated Opposition. *Folia Linguistica* 24, 4: 153-174.

Lotto, Andrew J. and Keith R. Kluender. 1998. General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics* 60, 4: 602-619.

Maddieson, Ian. 1995. Gestural Economy. *Proceedings of the International Conference on Phonetic Sciences, Stockholm* 4: 574-77.

Maddieson, Ian. 1984. *Patterns of sounds*. Cambridge (UK): Cambridge UP.

Martinet, André. 1955. *Economie des changements phonétiques*. Bern: Francke.

McCarthy, John J. 1991. Synchronic Rule Inversion. *Proceedings of the 17th Annual Meeting of the Berkley Linguistic Society* (L.A. Sutten, C. Johnson, and R. Shields, eds.). 192-207.

McDonough, Joyce and Keith Johnson. 1997. Tamil liquids: An investigation into the basis of the contrast among five liquids in a dialect of Tamil. *Journal of the International Phonetic Association* 27: 1-26.

McMahon, April M. S. 2000. *Lexical Phonology and the History of English*. Cambridge (UK): CUP.

Narayana, Shrikanth, Abeer Alwan and Kate Haker. 1995. An articulatory study of liquid approximants in American English. *Proceedings of the International Conference on Phonetic Sciences, Stockholm* 3: 576-79.

Newman, Paul. 1980. The two R's in Hausa. *African Language Studies* 17: 77-87.

Ni Chiosain, Maire and Jaye Padgett. Forthcoming. Markedness, segment realization, and locality in spreading. *Segmental phonology in Optimality Theory: constraints and representations* (L. Lombardi, editor). Cambridge (UK): CUP.

Nooteboom, S.G. and A. Cohen. 1995. *Spreken en verstaan: Een nieuwe inleiding tot de experimentele fonetiek*. Assen: Van Gorcum.

Ohala, John J. 1995. Phonetic explanations for sound patterns: Implications for grammars of competence. *Proceedings of the International Conference on Phonetic Sciences, Stockholm* 2: 52-59.

Padgett, Jaye. 2002. The Emergence of Contrastive Palatalization in Russian. http://www.ling.ecsc.edu/~padgett.

Padgett, Jaye. 2001. Contrast dispersion and Russian palatalization. *The role of speech perception in phonology* (E. V. Hume and K Johnson, editors). San Diego: Academic Press. 187-218.

Passy, Paul. 1891. *Etude sur les changements phonétiques et leurs caractères généraux*. Paris: Firmin-Didot.

Pétursson, Magnus. 1974. *Les articulations de l'islandais à la lumière de la radiocinématographie*. Strasbourg: Klincksieck.

Pétursson, Magnus. 1971. Etude de la réalisation des consonnes islandaises ɸ, ð, s, dans la prononciation d'un sujet islandais à partir de la radiocinématographie. *Phonetica* 23: 203-216.

Prokosch, Eduard. 1939. *A Comparative Germanic Grammar*. Philadelphia: LSA/U. Penn.

Recasens, Daniel. 1991. On the production characteristics of apicoalveolar taps and trills. *Journal of Phonetics* 19: 267-280.

Recasens, Daniel and Maria Dolors Pallarès. 1999. A study of /r/ and /ɾ/ in the light of the "DAC" coarticulation model. *Journal of Phonetics* 27: 143-169.

Reenen, Pieter van. 1994. Driemaal /r/ in Nederlandse dialecten. *Taal en Tongval* themanummer 7: 54-72.

Repp, Bruno H. 1988. Integration and Segregation in Speech Perception. *Language and Speech* 31, 3: 239-271.

Repp, Bruno H. 1986. The Role of Psychophysics in Understanding Speech Perception. *Haskins Laboratories Status Report on Speech Research* 86-87: 1-24.

Repp, Bruno H. and Alvin M. Liberman. 1984. Phonetic Category Boundaries are Flexible. *Haskins Laboratories Status Report on Speech Research* 77-78: 31-53.

Repp, Bruno H. and Virginia A. Mann. 1981. Fricative-Stop Coarticulation: Acoustic and Perceptual Evidence. *Haskins Laboratories Status Report on Speech Research* 67-68: 255-266.

Russell Webb, Eric. 2002. 'Dans la langue, il n'y a que des différences': what would consonant inventories look like if they were functionally optimal? To appear in *WECOL Review*.

Saussure, Ferdinand de. 1916/1995. *Cours de linguistique générale*. Paris: Payot.

Schiller, N.O. and C. Mooshammer. 1995. The character of /r/-sounds: Articulatory evidence for different reduction processes with special reference to German. *Proceedings of the International Conference on Phonetic Science, Stockholm* 3: 452-55.

Schutter, G. de and J. Taeldeman. 1994. Studies in r-klein. II. /r/ in hoofdtonige syllabe voor medeklinker. *Taal en Tongval* themanummer 7: 73-109.

Schwartz, Jean-Luc et al. 1997a. The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics* 25: 255-286.

Schwartz, Jean-Luc et al. 1997b. Major trends in vowel system inventories. *Journal of Phonetics* 25: 233-253.

Shadle, Christine H., Pierre Badin and André Moulinier. 1991. Towards the spectral characteristics of fricative consonants. *Actes du XIIème Congrès International des Sciences Phonétiques*, 19-24 August 1991: 42-45.

Shimizu, Katsumasa and Masatake Dantsuji. 1987. A Cross-Language Study on the Perception of [r-l]. *Studia Phonologica* 21: 10-18.

Sievers, Edward. 1898. *Angelsäsische Grammatik*, 3rd Edition. Halle: Niemeyer.

Spajic, Sinisa, Peter Ladefoged and Peri Bhaskararao. 1994. The rhotics of Toda. *UCLA Working Papers in Phonetics* 87: 35-66.

Stevens, Kenneth N. 1998. *Acoustic Phonetics*. Cambridge (MA): MIT Press.

Steriade, Donca. 1995. Underspecification and markedness. In *The Handbook of Phonological Theory* (John Goldsmith, ed.). Oxford: Blackwell. 114-174.

Steriade, Donca. 1990. Gestures and autosegments: comments on Browman and Goldstein's paper. *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech* (J. Kingston and M. Beckman, eds.). Cambridge (UK): CUP. 382-397.

Stevens, Kenneth N. 1989. On the quantal nature of speech. *Journal of Phonetics* 17: 3-45.

Straka, Georges. 1979. *Les sons et les mots: Choix d'études de phonétique et de linguistique*. Strasbourg: Klincksieck.

Tobin, Yishai. 1997. *Phonology as Human Behavior: Theoretical Implications and Clinical Applications*. Durham (NC): Duke UP.

Tousignant, Claude. 1987. Les variantes du /ʀ/ montréalais: contextes phonologiques favorisant leur apparition. *Revue québecoise de linguistique théorique et appliquée* 6, 3: 73-113.

Tranel, Bernard. 1987. *The sounds of French*. Cambridge (UK): Cambridge UP.

Trautmann, Moritz. 1886. *Die Sprachlaute: Im allgemeinen, und die Laute des Englischen, Franzosischen und Deutschen im besondern*. Leipzig.

Trubetzkoy, N.S. 1939. *Grundzüge der Phonologie*. Prague: Travaux du Cercle Linguistique de Prague.

Uldall, Elizabeth. 1958. American "Molar" R and "Flapped" T. *Revista do laboratorio de Fonetica Experimental* 4: 103-106.

Velde, Hans van de. 1994. 60 jaar (r)evolutie in het Standaard-Nederlands. *Taal en Tongval* themanummer 7: 22-42.

Walsh Dickey, Laura. 1997. *The Phonology of Liquids*. Ph.D. Dissertation, University of Massachusetts at Amherst.

Wängler, Hans-Heinrich. 1976. *Atlas Deutscher Sprachlaute*. Berlin: Akademie Verlag.

Werlen, Iwar. 1980. R im Schweizerdeutschen. *Zeitschrift fur Dialektologie und Linguistik* 47: 52-76.

West, Paula. 1999. Perception of distributed coarticulatory properties of English /l/ and /r/. *Journal of Phonetics* 27: 405-426.

Westbury, John R., Michiko Hashi, and Mary J. Lindstrom. 1998. Differences among speakers in lingual articulations for American English /r/. *Speech Communication* 26: 203-226.

Widdison, Kirk A. 1997. Variability in lingual vibrants: Changes in the story of /r/. *Language and Communication* 17, 3: 187-193.

Zipf, George. 1949. Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology. Cambridge (MA): MIT Press.

VITA

Eric Russell Webb was born in Heidelberg/N. (Germany) on 7 May 1973, the son of Caroline Russell (Parker). He studied at l'Université de Strasbourg II—Sciences Humaines, and there completed a *DEUG* in French philology in 1994. He received a Bachelor of Arts, magna cum laude, in French and Linguistics at Houghton College, Houghton, New York in 1995. He completed a *licence* in linguistics at l'Université François Rabelais – Tours in 1996 and a Master of Arts in French Studies with a concentration in sociolinguistics at Bowling Green State University in 1997. He entered the Graduate School of the University of Texas at Austin in 1997.

Permanent address: 7 Russell Farm Road; Kennebunk, Maine 04043 USA

This dissertation was typed by the author.